Figure 1

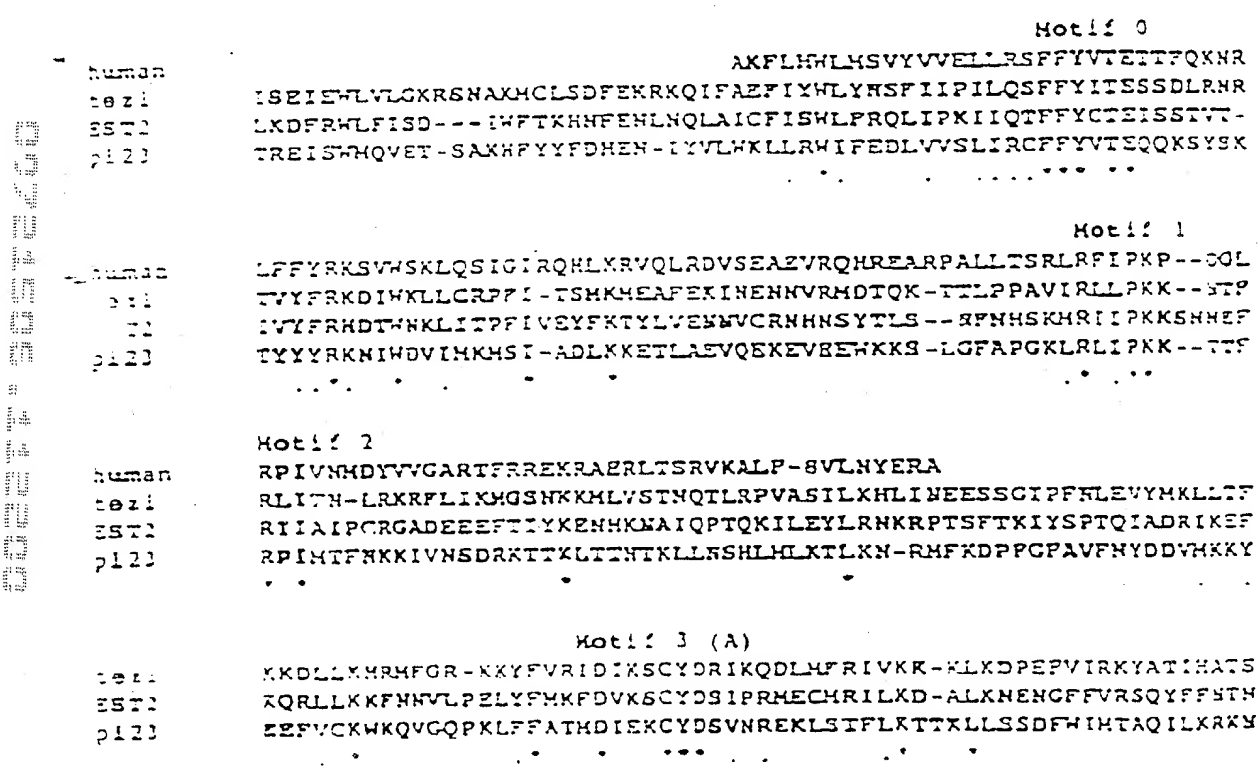

Motif 0
human
tezi
ESTi
p123
Motif 1
human
tezi
ESTi
p123
Motif 2
human
tezi
ESTi
p123
Motif 3 (A)
tezi
ESTi
p123

Figure 2

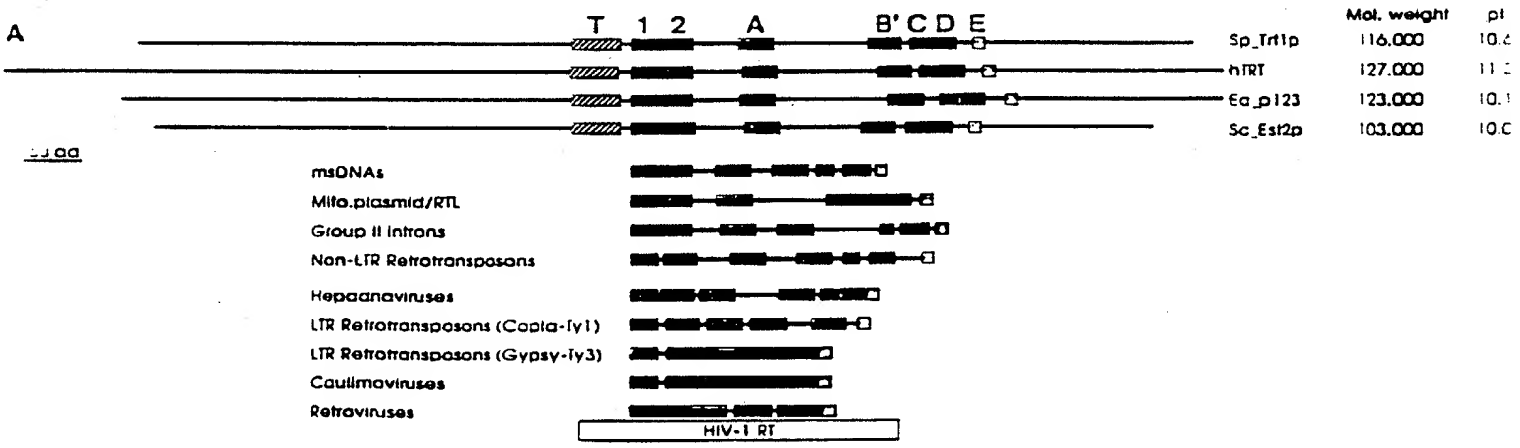

A
T 1 2 A B' C D E
Mol. weight
pI
Sp_Trt1p
116.000
hTRT
127.000
Ea_p123
123.000
Sc_Est2p
103.000
msDNAs
Mito.plasmid/RTL
Group II Introns
Non-LTR Retrotransposons
Hepadnaviruses
LTR Retrotransposons (Copia-Ty1)
LTR Retrotransposons (Gypsy-Ty3)
Caulimoviruses
Retroviruses
HIV-1 RT

Figure 3

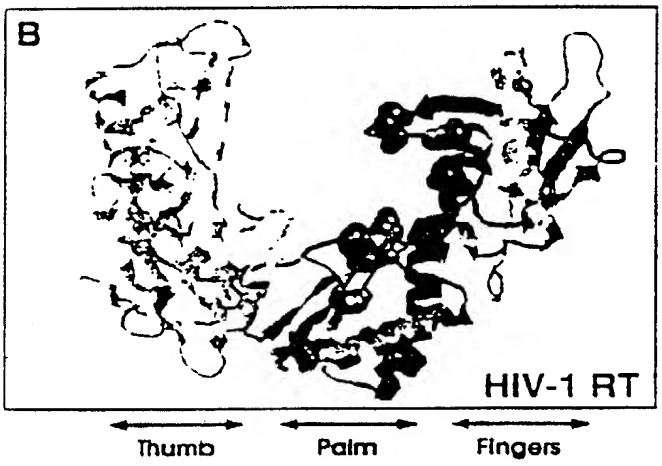


Figure 4

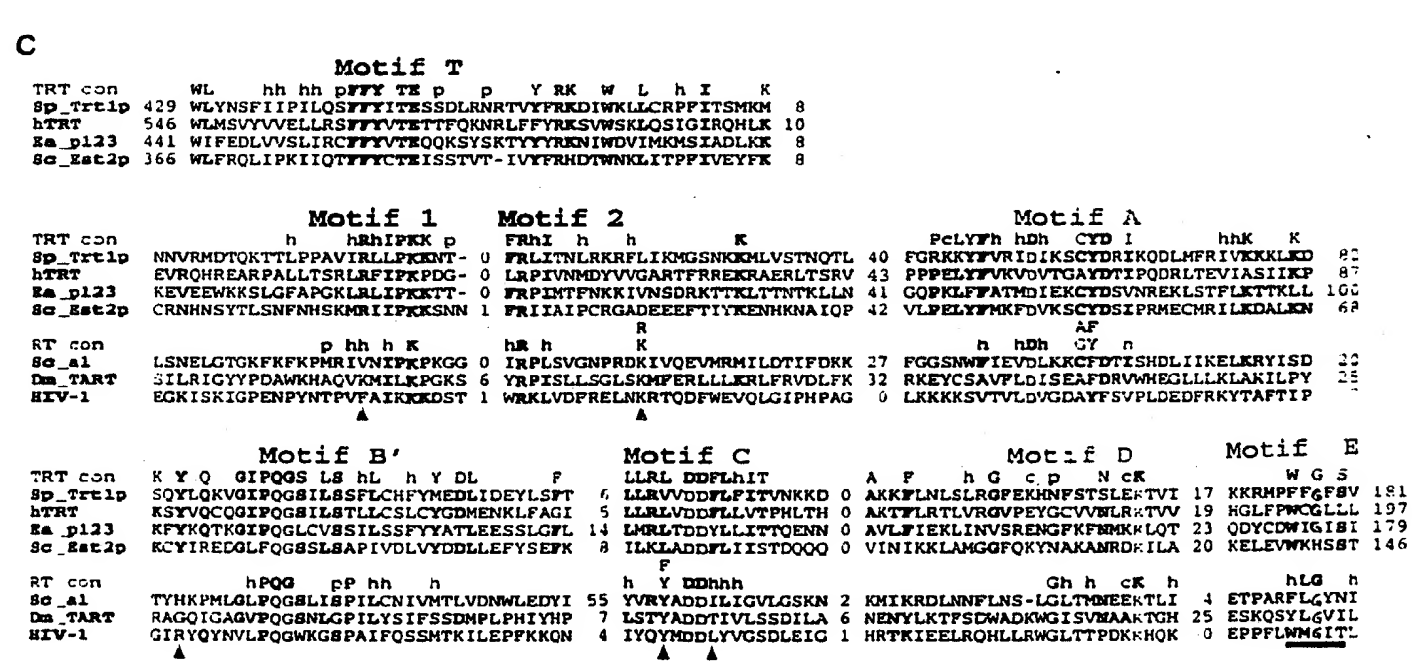

C
Motif T
TRT con WL hh hh pFFY TE p p Y RK W L h I K
Sp_Trt1p 429 WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFITSMKM 8
hTRT 546 WLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLK 10
Ka_p123 441 WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSIADLKK 8
Sc_Est2p 366 WLFRQLIPKIIQTFFYCTEISSTVT-IVYFRHDTWNKLITPFIVEYFK 8
Motif 1
Motif 2
Motif A
TRT con h hRhIPKK p FRhI h h K PcLYFh hDh CYD I hhK K
Sp_Trt1p NNVRMDTQKTTLPPAVIRLLPKKNT- 0 FRLITNLRKRFLIKMGSNKKMLVSTNQTL 40 FGRKKYFVRIDIKSCYDRIKQDLMFRIVKKKLKD 8
hTRT EVRQHREARPALLTSRLRFIPKPDG- 0 LRPIVNMDYVVGARTFRREKRAERLTSRV 43 PPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKP 8
Ka_p123 KEVEEWKKSLGFAPGKLRLIPKKTT- 0 FRPIMTFNKKIVNSDRKTTKLTTNTKLLN 41 GQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLL 10
Sc_Est2p CRNHNSYTLSNFNHSKMRIIPKKSNN 1 FRIIAIPCRGADEEEFTIYKENHKNAIQP 42 VLPELYFMKFDVKSCYDSIPRMECMRILKDALKN 6
RT con p hh h K hR h K h hDh GY n
Sc_a1 LSNELGTGKFKFKPMRIVNIPKPKGG 0 IRPLSVGNPRDKIVQEVMRMILDTIFDKK 27 FGGSNWFIEVDLKKCFDTISHDLIIKELKRYISD 3
Dm_TART SILRIGYYPDAWKHAQVKMILKPGKS 6 YRPISLLSGLSKMFERLLLKRLFRVDLFK 32 RKEYCSAVFLDISEAFDRVWHEGLLLKLAKILPY 2
HIV-1 EGKISKIGPENPYNTPVFAIKKKDST 1 WRKLVDFRELNKRTQDFWEVQLGIPHPAG 0 LKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIP
Motif B'
Motif C
Motif D
Motif E
TRT con K Y Q GIPQGS LS hL h Y DL F LLRL DDFLhIT A F h G c p N cK W G S
Sp_Trt1p SQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFT 6 LLRVVDDFLFITVNKKD 0 AKKFLNLSLRGFEKHNFSTSLEFTVI 17 KKRMPFFGFSV 181
hTRT KSYVQCQGIPQGSILSTLLCSLCYGDMENKLFAGI 5 LLRLVDDFLLVTPHLTH 0 AKTFLRTLVRGVPEYGCVVNLRKTVV 19 HGLFPWCGLLL 197
Ka_p123 KFYKQTKGIPQGLCVSSILSSFYYATLEESSLGFL 14 LMRLTDDYLLITTQENN 0 AVLFIEKLINVSRENGFKFNMKKLQT 23 QDYCDWIGISI 179
Sc_Est2p KCYIREDGLFQGSSLSAPIVDLVYDDLLEFYSEFK 8 ILKLADDFLIISTDQQQ 0 VINIKKLAMGGFQKYNAKANRDRILA 20 KELEVWKHSST 146
RT con hPQG pP hh h h Y DDhhh Gh h cK h hLG h
Sc_a1 TYHKPMLGLPQGSLISPILCNIVMTLVDNWLEDYI 55 YVRYADDILIGVLGSKN 2 KMIKRDLNNFLNS-LGLTMNEEKTLI 4 ETPARFLGYNI
Dm_TART RAGQIGAGVPQGSNLGPILYSIFSSDMPLPHIYHP 7 LSTYADDTIVLSSDILA 6 NENYLKTFSDWADKWGISVNAAKTGH 25 ESKQSYLGVIL
HIV-1 GIRYQYNVLPQGWKGSPAIFQSSMTKILEPFKKQN 4 IYQYMDDLYVGSDLEIG 1 HRTKIEELRQHLLRWGLTTPDKKHQK 0 EPPFLWMGITL

Figure 5

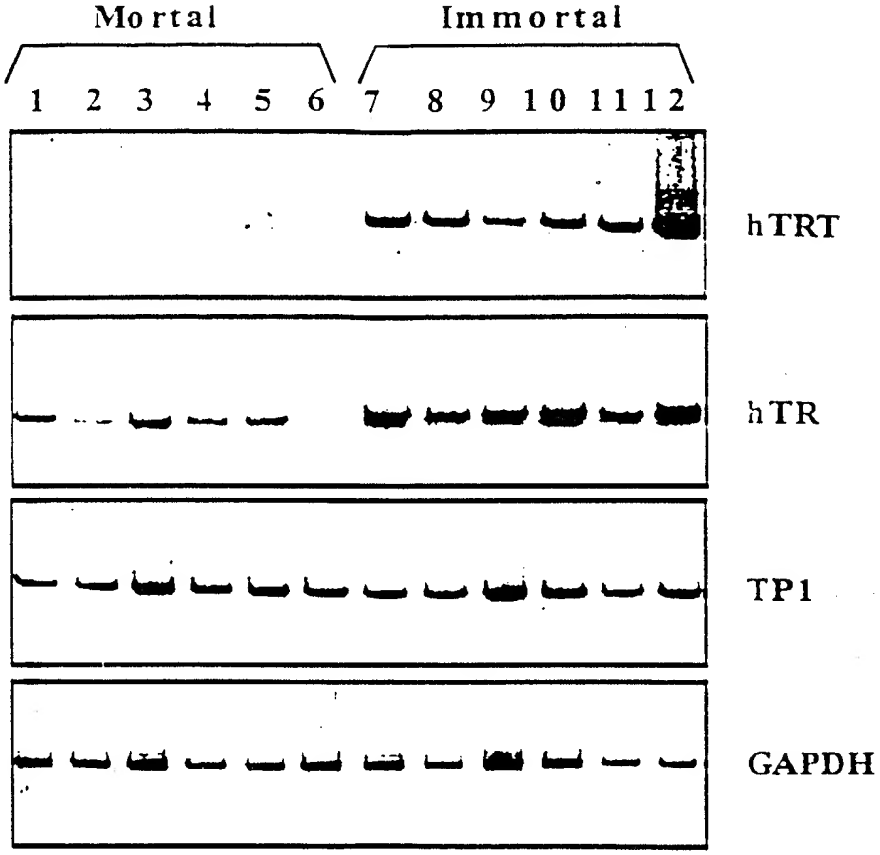


Figure 6

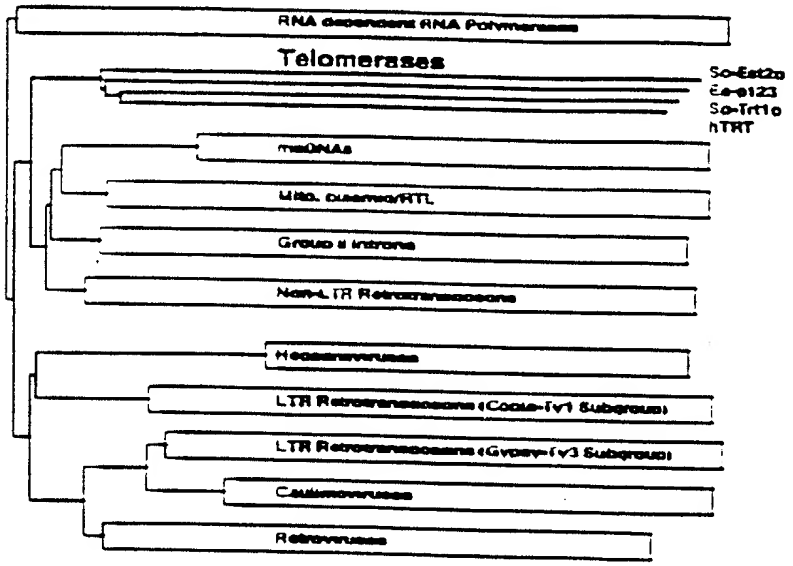

RNA dependent RNA Polymerases
Telomerases
So-Est2p
Ee-e123
So-Trt1p
hTRT
mtSINAs
Group II introns
Non-LTR Retrotransposons
Hepadnaviruses
LTR Retrotransposons (Copia-like Subgroup)
LTR Retrotransposons (Gypsy-like Subgroup)
Caulimoviruses
Retroviruses

Figure 7

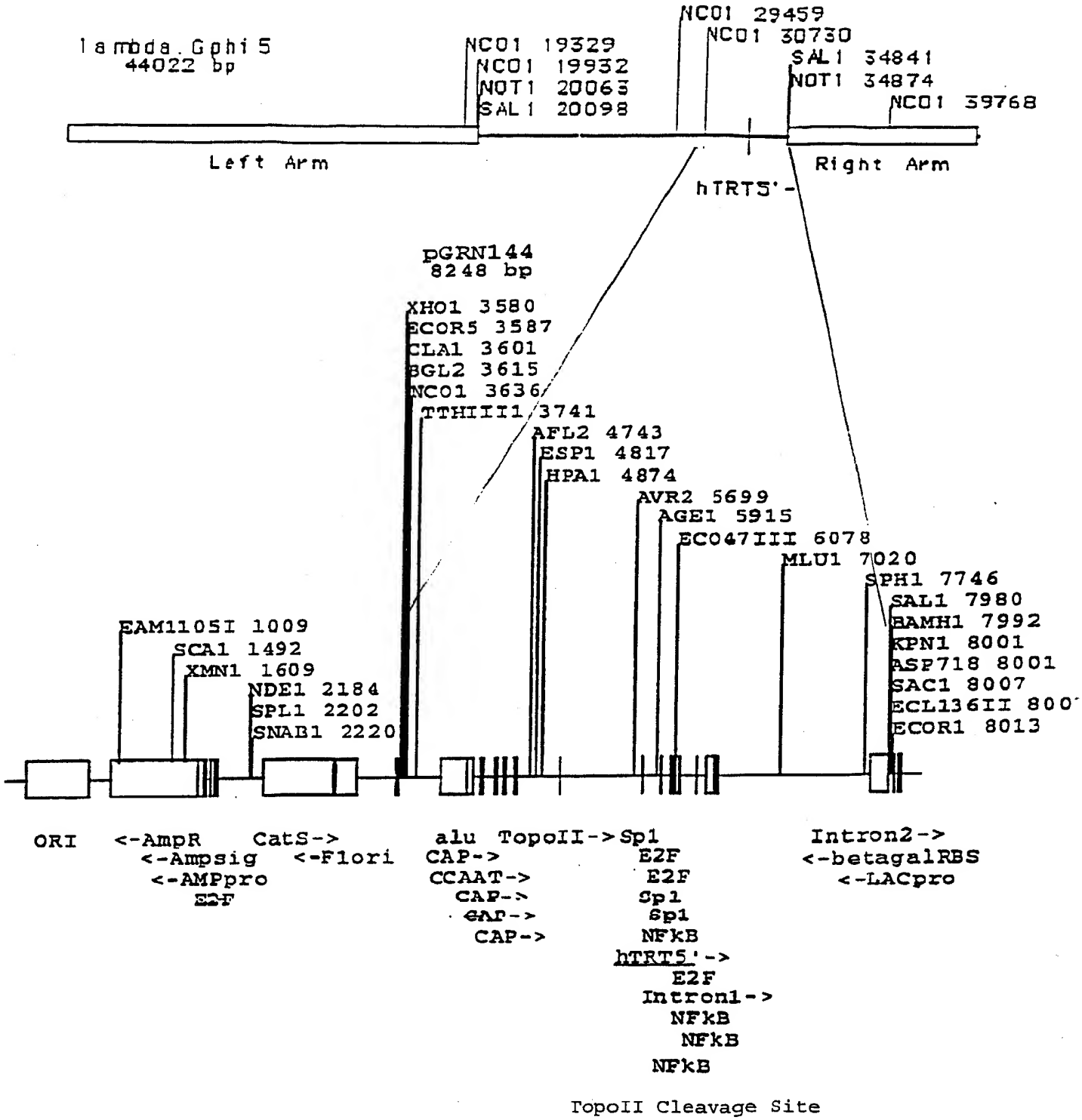

lambda.Gphi5
44022 bp
NCO1 19329
NCO1 19932
NOT1 20063
SAL1 20098
NCO1 29459
NCO1 30730
SAL1 34841
NOT1 34874
NCO1 39768
Left Arm
Right Arm
hTRT5'-
pGRN144
8248 bp
XHO1 3580
ECOR5 3587
CLA1 3601
BGL2 3615
NCO1 3636
TTHIII1 3741
AFL2 4743
ESP1 4817
HPA1 4874
AVR2 5699
AGE1 5915
EC047III 6078
MLU1 7020
SPH1 7746
SAL1 7980
BAMH1 7992
KPN1 8001
ASP718 8001
SAC1 8007
ECL136II 800
ECOR1 8013
EAM1105I 1009
SCA1 1492
XMN1 1609
NDE1 2184
SPL1 2202
SNAB1 2220
ORI
<-AmpR
<-Ampsig
<-AMPpro
E2F
CatS->
<-Flori
alu
CAP->
CCAAT->
CAP->
CAP->
TopoII->
Sp1
E2F
E2F
Sp1
Sp1
NFkB
hTRT5'->
E2F
Intron1->
NFkB
NFkB
NFkB
Intron2->
<-betagalRBS
<-LACpro
TopoII Cleavage Site

Figure 8

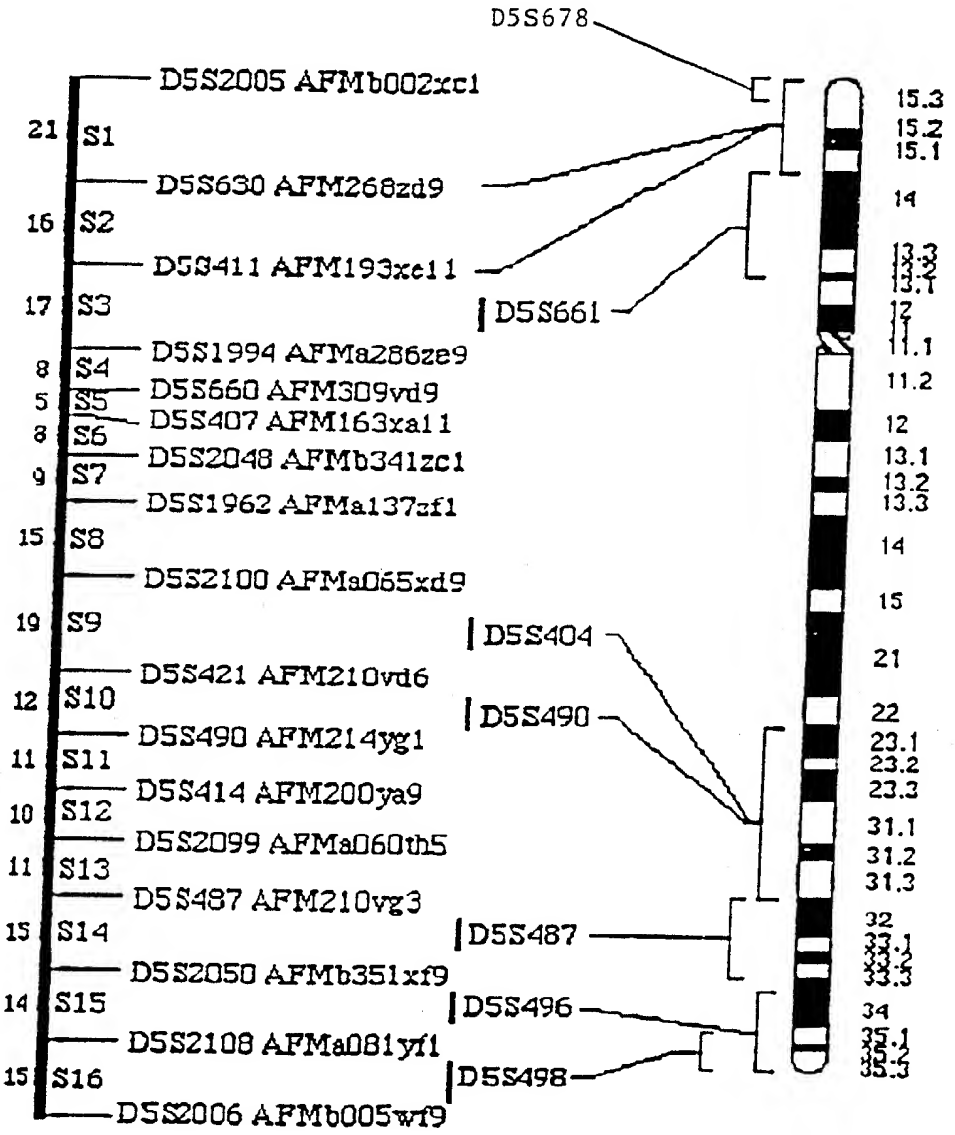

D5S678
D5S2005 AFMb002xc1
21 S1
D5S630 AFM268zd9
16 S2
D5S411 AFM193xe11
17 S3
D5S661
D5S1994 AFMa286ze9
8 S4
D5S660 AFM309vd9
5 S5
D5S407 AFM163xa11
8 S6
D5S2048 AFMb341zc1
9 S7
D5S1962 AFMa137zf1
15 S8
D5S2100 AFMa065xd9
19 S9
D5S404
D5S421 AFM210vd6
12 S10
D5S490
D5S490 AFM214yg1
11 S11
D5S414 AFM200ya9
10 S12
D5S2099 AFMa060th5
11 S13
D5S487 AFM210vg3
15 S14
D5S487
D5S2050 AFMb351xf9
14 S15
D5S496
D5S2108 AFMa081yf1
15 S16
D5S498
D5S2006 AFMb005wf9
15.3
15.2
15.1
14
13.3
13.2
13.1
12
11
11.1
11.2
12
13.1
13.2
13.3
14
15
21
22
23.1
23.2
23.3
31.1
31.2
31.3
32
33.1
33.2
33.3
34
35.1
35.2
35.3

FIGURE 9

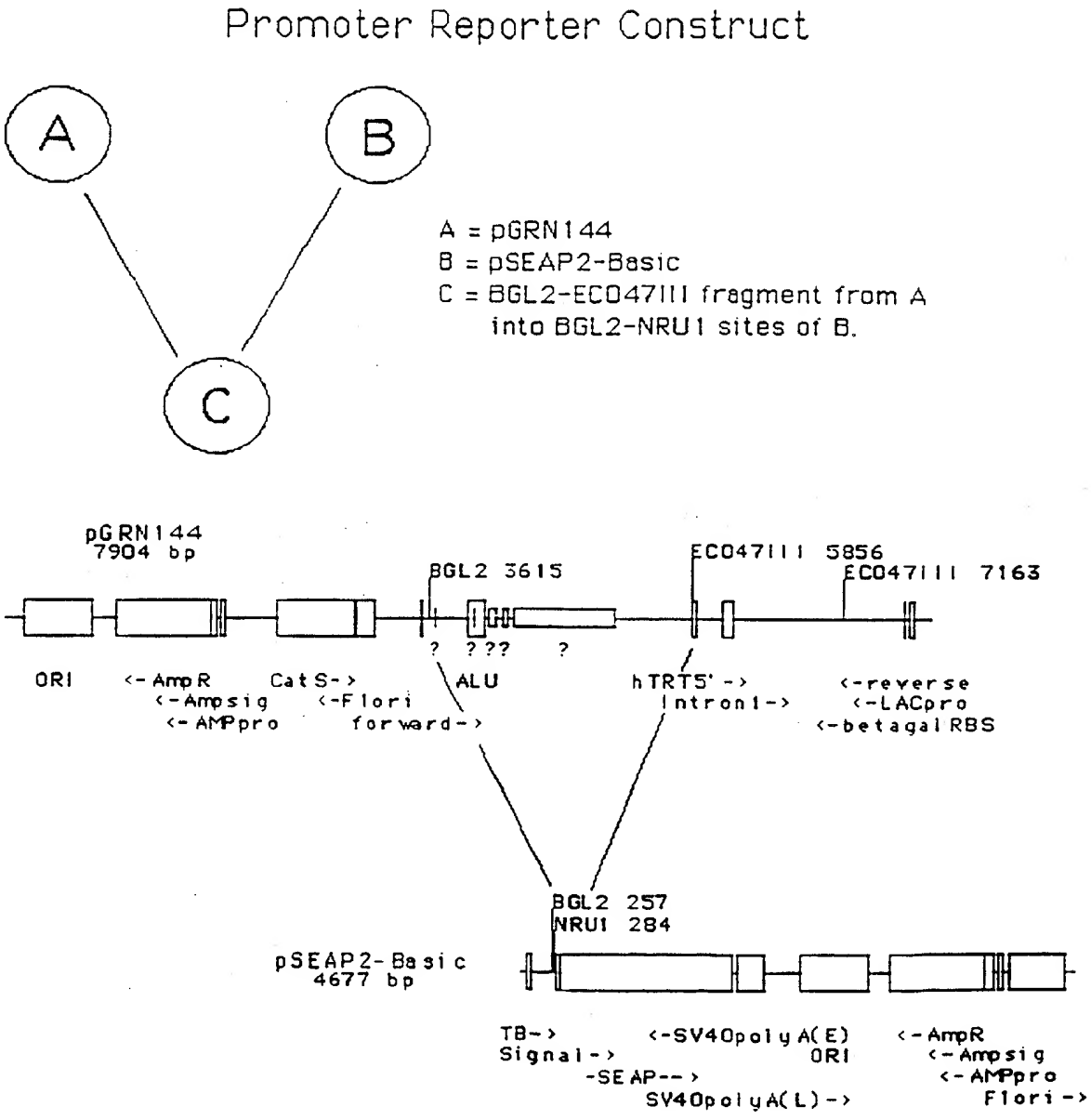
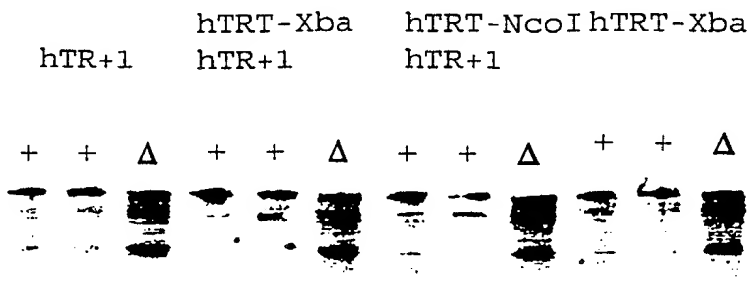

Promoter Reporter Construct
A
B
C
A = pGRN144
B = pSEAP2-Basic
C = BGL2-EC047III fragment from A
into BGL2-NRU1 sites of B.
pGRN144
7904 bp
BGL2 3615
EC047III 5856
EC047III 7163
? ??? ?
ORI
<-AmpR
<-Ampsig
<-AMPpro
CatS->
<-Flori
forward->
ALU
hTRT5'->
Intron1->
<-reverse
<-LACpro
<-betagalRBS
BGL2 257
NRU1 284
pSEAP2-Basic
4677 bp
TB->
Signal->
-SEAP-->
<-SV40polyA(E)
ORI
SV40polyA(L)->
<-AmpR
<-Ampsig
<-AMPpro
Flori->
pSEAP2-Basic

Figure 10
Page 1



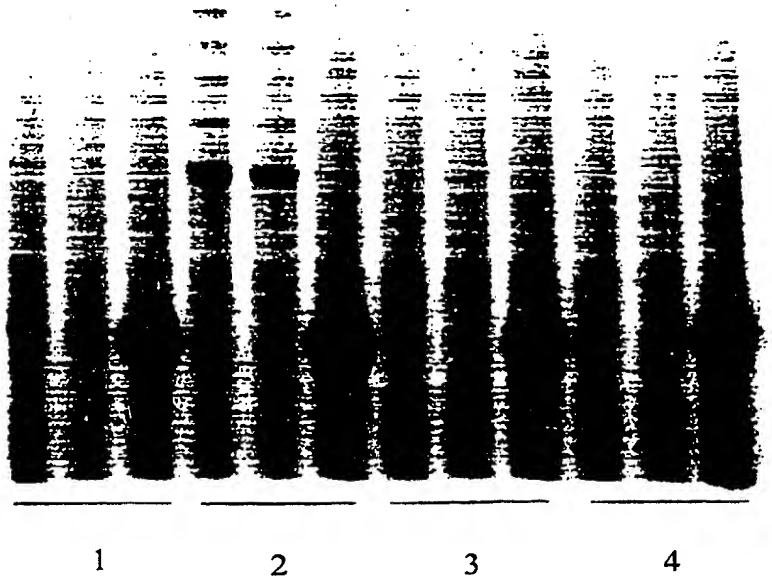


Figure 10

Page 2

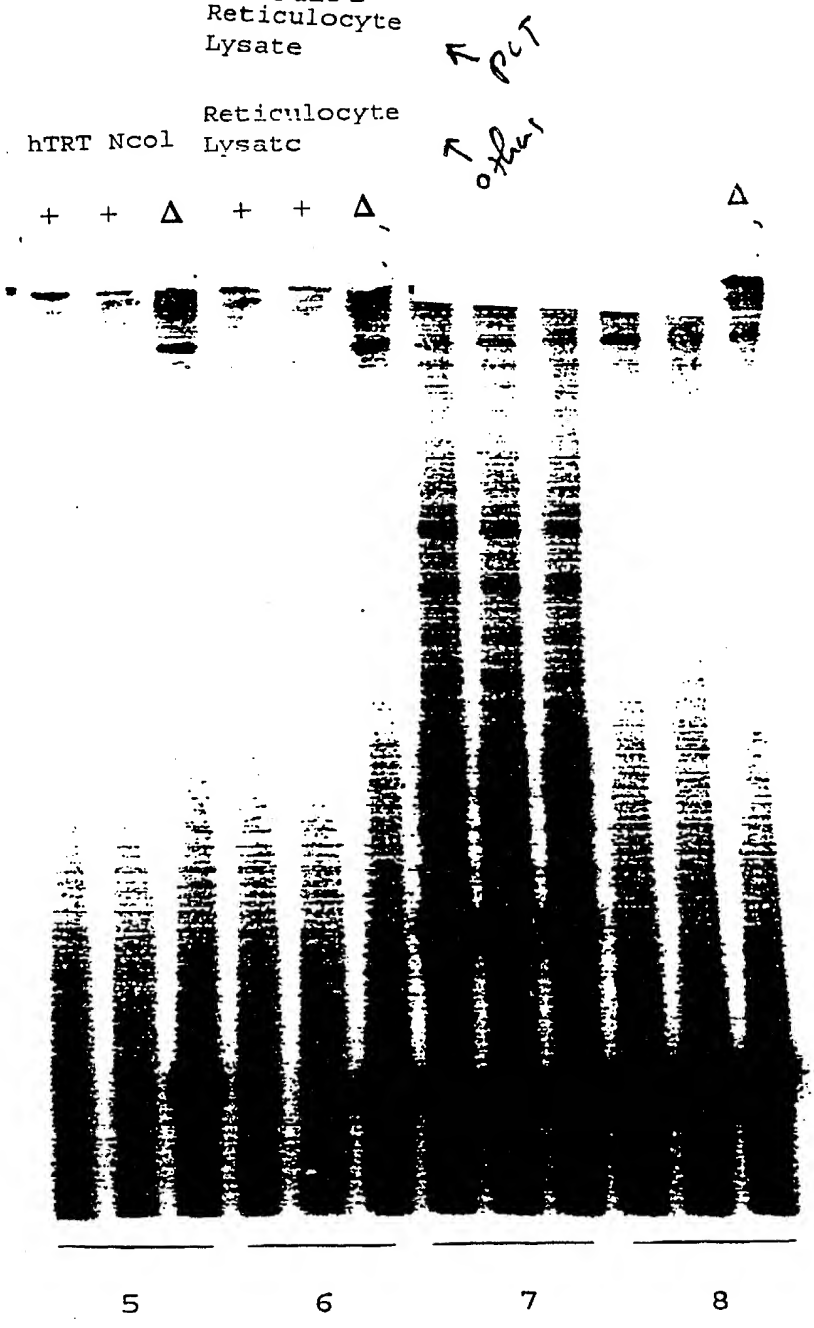


FIGURE 11
Page 1

Telomerase Specific Motifs

```
                          MOTIF T                                    MOTIF T'
TRT con        Wl             FFY TE          y Rk  W  l    I     E  V
hTRT      546 WLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGI 13 EAEVR
spTRT     429 WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFI 12 ENNVR
Ea_p123   441 WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSI 12 EKEVE
Sc_Est2   366 WLFRQLIPKIIQTFFYCTEISSTVT.IVYFRHDTWNKLITPFI  9 ENNVC
```

Telomerase RT Motifs (Fingers)

```
              MOTIF 1    MOTIF 2          MOTIF A              MOTIF B'
TRT con         R iPKk     fR I        p lyF   D   cYD i     Y q  GipQGs lS  l    y
hTRT      11 SRLRFIPKPDG 0 LRPIV    69 PELYFVKVDVTGAYDTI 104 YVQCQGIPQGSILSTLLCSLCY
spTRT     10 AVIRLLPKKNT 0 FRLIT    66 RKKYFVRIDIKSCYDRI  99 YLQKVGIPQGSILSSFLCHFYM
Ea_p123   10 GKLRLIPKKTT 0 FRPIM    67 PKLFFATMDIEKCYDSV 117 YKQTKGIPQGLCVSSILSSFYY
Sc_Est2   13 SKMRIIPKKSN 2 FRIIA    68 PELYFMKFDVKSCYDSI  85 YIREDGLFQGSSLSAPIVDLVY
RT con         p hh h K     hR h          h  hDh  AF  h          hPQG   pP hh   h
                                          GY
```

FIGURE 11
Page 2

Telomerase RT Motifs (Palm, Primer Grip)

	MOTIF C		MOTIF D		MOTIF E	
TRT con	lllrl DDfL it		g n K		w g s l	
hTRT	15 LLLRLVDDFLLVT	15	GVPEYGCVVNLRKTVV	24	WCGLLLDTRTL	192
spTRT	16 VLLRVVDDFLFIT	15	GFEKHNFSTSLEKTVI	22	FFGFSVNMRSL	176
Ea_p123	24 LLMRLTDDYLLIT	15	VSRENGFKFNMKKLQT	28	WIGISIDMKTL	174
Sc_Est2	18 LILKLADDFLIIS	15	GFQKYNAKANRDKILA	25	WKHSSTMNNFH	141
RT con	h Y DDhhh F		Gh h cK h		hLG h	

FIGURE 12

Seq. ID. No 7

```
181 GGACCCGGCGGCTTTCCGCGCGCTGGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACGC
    CCTGGGCCGCCGAAAGGCGCGCGACCACCGGGTCACGGACCACACGCACGGGACCCTGCG

                                                 NFkB_CS1
                                                 GGGRQTYYQC
                                                NFkB-MHC-I.2
                                                TGGGCTTCCCC
                                               **************************
241 ACGGCCGCCCCCCGCCGCCCCCTCCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCG
    TGCCGGCGGGGGGCGGCGGGGGAGGAAGGCGGTCCACCCGGAGGGGCCCCAGCCGCAGGC

                              Intron1
    ************************************************************
301 GCTGGGGTTGAGGGCGGCCGGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTC
    CGACCCCAACTCCCGCCGGCCCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAG

       NFkB_CS1
      GGGRQTYYQC
       NFkB_CS2
     RGGGRMTYYCC
        Topo_II_cleavage_site
          RNYNNCNNGYNGKTNYNY
    ******************>
361 AGGGCGCTTCCCCCGCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGG
    TCCCGCGAAGGGGGCGTCCACAGGACGGACTTCCTCGACCACCGGGCTCACGACGTCTCC
```

Figure 13


```
   1 AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT
  51 AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA
 101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC
 151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA
 201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
 251 TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG
 301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT
 351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACTTCTT
 401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTCA
 451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTCAA
 501 CTCAAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
 551 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA
 601 GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG
 651 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA
 701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
 751 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA
 801 TATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT TAATAGAAAT
 851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTC
 901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA
 951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG
1001 CTTGAGAAAG TCAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG
1051 TCCTCTTCCA GAAAATTGGC GGGAACGGAA ACAAAAAAATC GAAAACTTGA
1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC
1151 TACACAACTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA
1201 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC
1251 AAAAGAAAGT TAAGAAATAT GTGGAACTAA ACAAGCATGA ACTCATTCAC
1301 AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
1351 GGTTGAGACC TCTGCAAAGC ATTTTTATTA TTTTGATCAC GAAAACATCT
1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
1451 CTGATTAGAT GATTTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
1601 GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAAC TCAGACTAAT
1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
1701 TAAATTCAGA CCGGAAGACT ACAAAATTAA CTACAAATAC GAAGTTATTG
1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
1801 TTTTGGATTC GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
1851 AGTTTGTTTG CAAATGGAAG CAAGTTGGAC AACCAAAACT CTTCTTTGCA
1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAACTATC
1951 AACATTCCTA AAAACTACTA AATTACTTTC TTCAGATTTC TGGATTATGA
2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC
2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
2101 TGCACTTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG
2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTTG
2251 CCAATATAAT TACATTAACT TTAATGGGAA GTTTTATAAA CAAACAAAAG
2301 GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTCATC ATTTTATTAT
```

Figure 13
Page 2

2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA
2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
2551 GACTAGTTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
2651 TCAATTGATA TGAAAACTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC
2801 CATTATTTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAAACTCT
2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
2901 AATACAAGGA CCACTTTAAG AAGAACTTAG CTATGAGCAG TATGATCGAC
2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
3001 CCTTGTGTGC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG
3051 ACTTTTTCCT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA
3101 AAGTACATTT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA
3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
3201 CTATTCTAAC TTATTTTGGA AAGTTAATTT TCAATTTTTG TCTTATATAC
3251 TGGGGTTTTG GGGTTTTGGG GTTTTGGGG

Figure 14

```
   1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD
  51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
 101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
 151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGA
 201 ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNNMKSRTR IFYCTHFNRN
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAYM
 301 LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHELIH
 401 KNLLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKLL RWIFEDLVVS
 451 LIRCFFYVTE QQKSYSKTYY YRKNIWDVIM KMSIADLKKE TLAEVQEKEV
 501 EEWKKSLGFA PGKLRLIPKK TTFRPIMTFN KKIVNSDRKT TKLTTNTKLL
 551 NSHLMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN
 651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLIVEAKQ
 701 RNYFKKDNLL QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFYY
 751 ATLEESSLGF LRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIEKL
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
 851 SIDMKTLALM PNINLRIEGI LCTLNLNMQT KKASMWLKKK LKSFLMNNIT
 901 HYFRKTITTE DFANKTLNKL FISGGYKYMQ CAKEYKDHFK KNLAMSSMID
 951 LEVSKIIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK
1001 KYLFNRVCMI LKAKEAKLKS DQCQSLIQYD A
```

Figure 15


```
1 ggtaccgatttacttttcttttcttcataagctaattgcttcctcgaacgctcctaaatcttctggaaatatttttacaaga 80
81 acttaataacaatacctaagtcaaattccaatatgaaggtgttattagtgatcgataatatttctatttttatcggtcgtta 160
161 ccaagtataaggacaaaaagaacaacttccttcccccctaaagacttttactttattaatttacttttcaaatatatttccg 240
241 ggttcgcttacttttaatcgtggtactgttttagctgctacttctagccaaccgcgtgtttctaccccgttcattggatat 320
321 agctcttggagtagctcacagaaatccttacaaatcttctgatgagactatattagattcattacagtccgtgcatattc 400
401 ttaacatggagccttacactttagatgagtcacgtcgcatgatggagtatttggtatcatccaacgtttgccttgaaaag 480
481 gttgataattatttgcaaaatcatgtccttagtggtggtaatccgcgaaagtttttgatgcttgcacacgtcttagcatg 560
561 attgagatattcaaaaatttctatccacttacaactcctttaacgcggttttatttttctatttttctattcttcatgttgtt 640
641 ccaaatatgtatcatctcgtattaggcttttttccgttttacttcctggaatcgtaccttttttcactattcccccctaatga 720
721 ataatctaaattagtttcgcttataattgatagtagtagaaagattggtgattcttacttcgtgtaatgttattagtttaaa 800
801 gatacttttgcaaaacatttattagcttatcattatataaaaaaaaatccctataattataaatattaatcaatatttgcggtc 880
881 acttatttatttaaaacgttatgatcagtaggacactttgcatatatatagttatgcttaatggttacttgtaacttgc 958

959 ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018
  1 M   T   E   H   H   T   P   K   S   R   I   L   R   F   L   E   N   Q   Y   V  20

1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078
  21 Y   L   C   T   L   N   D   Y   V   Q   L   V   L   R   G   S   P   A   S   S  40

1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138
  41 Y   S   N   I   C   E   R   L   R   S   D   V   Q   T   S   F   S   I   F   L  60

1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198
  61 H   S   T   V   V   G   F   D   S   K   P   D   E   G   V   Q   F   S   S   P  80
```

Figure 15
Page 2

```
1199 AAA TGC TCA CAG TCA GAG gtatatatattttgtttgatttttttctattcgggatagctaatatatgggcag 1272
  81 K   C   S   Q   S   E                                                          86

1273 CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA 1332
  87 L   I   A   N   V   V   K   Q   M   F   D   E   S   F   E   R   R   R   N   L  106

1333 CTG ATG AAA GGG TTT TCC ATG gtaaggtatttctaattgtgaaatatttacctgcaacttactgtttcaaagaga 1405
 107 L   M   K   G   F   S   M                                                      113

1406 ttgtatttaaccgataaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT 1469
 114                     N   H   E   D   F   R   A   M   H   V   N   G   V   Q   N  128

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA 1529
 129 D   L   V   S   T   F   P   N   Y   L   I   S   I   L   E   S   K   N   W   Q  148

1530 CTT TTG TTA GAA AT gtaaataccggttaagatgttgcgcactttgaacaagactgacaagtatag T ATC GGC 1601
 149 L   L   L   E   I                                                          I   G  155

1602 AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661
 156 S   D   A   M   H   Y   L   L   S   K   G   S   I   F   E   A   L   P   N   D  175

1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721
 176 N   Y   L   Q   I   S   G   I   P   L   F   K   N   N   V   F   E   E   T   V  195

1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA 1781
 196 S   K   K   R   K   R   T   I   E   T   S   I   T   Q   N   K   S   A   R   K  215

1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT 1841
 216 E   V   S   W   N   S   I   S   I   S   R   F   S   I   F   Y   R   S   S   Y  235

1842 AAG AAG TTT AAG CAA G gtaactaatactgttatcctttataactaatttag AT CTA TAT TTT AAC
1907
 236 K   K   F   K   Q   D                                                      L   Y   F   N  245

1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG 1967
 246 L   H   S   I   C   D   R   N   T   V   H   M   W   L   Q   W   I   F   P   R  265

1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA 2027
 266 Q   F   G   L   I   N   A   F   Q   V   K   Q   L   H   K   V   I   P   L   V  285

2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087
 286 S   Q   S   T   V   V   P   K   R   L   L   K   V   Y   P   L   I   E   Q   T  305

2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT 2147
 306 A   K   R   L   H   R   I   S   L   S   K   V   Y   N   H   Y   C   P   Y   I  325

2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207
 326 D   T   H   D   D   E   K   I   L   S   Y   S   L   K   P   N   Q   V   F   A  345

2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267
 346 F   L   R   S   I   L   V   R   V   F   P   K   L   I   W   G   N   Q   R   I  365

2268 TTT GAG ATA ATA TTA AAA G gtattgtataaaacttattaccacttaacgattttaccag AC TTC GAA ACT 2336
 366 F   E   I   I   L   K   D                                                  L   E   T  375
```

Figure 15


```
2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396
 376 F   L   K   L   S   R   Y   E   S   F   S   L   H   Y   L   M   S   N   I   K  395

2397 gtaatatgccaaatttttttaccattaattaacaatcag ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465
 396                                        I   S   E   I   E   W   L   V   L   G  405

2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525
 406 K   R   S   N   A   K   M   C   L   S   D   F   E   K   R   K   Q   I   F   A  425

2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585
 426 E   F   I   Y   W   L   Y   N   S   F   I   I   P   I   L   Q   S   F   F   Y  445

2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645
 446 I   T   E   S   S   D   L   R   N   R   T   V   Y   F   R   K   D   I   W   K  465

2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705
 466 L   L   C   R   P   F   I   T   S   M   K   M   E   A   F   E   K   I   N   E  485

2706 gtattttaaagtatttttttgcaaaaagctaatattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775
 486                                          N   N   V   R   M   D   T   Q   K   T  495

2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835
 496 T   L   P   P   A   V   I   R   L   L   P   K   K   N   T   F   R   L   I   T  515

2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtattaatttttggtcatcaatgtactttactttttaatctatta 2906
 516 N   L   R   K   R   F   L   I   K                                               524

2907 ttagcag ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967
 525       M   G   S   N   K   K   M   L   V   S   T   N   Q   T   L   R   P   V  542

2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027
 543 A   S   I   L   K   H   L   I   N   E   E   S   S   G   I   P   F   N   L   E  562

3028 GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT CTT AAG CAC CGA ATG TTT GG gtaat 3088
 563 V   Y   M   K   L   L   T   F   K   K   D   L   L   K   H   R   M   F   G      581

3089 tatataatgcgcgattcctcattattaattttgcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA 3155
 582                                     R   K   K   Y   F   V   R   I   D   I     591

3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC 3215
 592 K   S   C   Y   D   R   I   K   Q   D   L   M   F   R   I   V   K   K   K   L  611

3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT 3275
 612 K   D   P   E   F   V   I   R   K   Y   A   T   I   H   A   T   S   D   R   A  631

3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagtttatttttcattggaatttttaacaa 3343
 632 T   K   N   F   V   S   E   A   F   S   Y   F                                  643

3344 atttttttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA 3405
 644            D   M   V   P   F   E   K   V   V   Q   L   L   S   M   K   T  659

3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT 3465
 660 S   D   T   L   F   V   D   F   V   D   Y   W   T   K   S   S   S   E   I   F  679

3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaatcgttgaattgtaataaca 3532
 680 K   M   L   K   E   H   L   S   G   H   I   V   K                              692
```

Figure 15
Page 4

```
3533 ctaatgaaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA   3593
693                I   G   N   S   Q   Y   L   Q   K   V   G   I   P   Q   G   S    708

3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG 3653
709  I   L   S   S   F   L   C   H   F   Y   M   E   D   L   I   D   E   Y   L   S   728

3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA 3713
729  F   T   K   K   K   G   S   V   L   L   R   V   V   D   D   F   L   F   I   T   748

3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgctgtcattcc 3777
749  V   N   K   K   D   A   K   K   F   L   N   L   S   L   R   G                    764

3778 caagttctaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA 3840
765                        F   E   K   H   N   F   S   T   S   L   E   K   T   V    778

3841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA 3900
779  I   N   F   E   N   S   N   G   I   I   N   N   T   F   F   N   E   S   K   K   798

3901 AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT 3960
799  R   M   P   F   F   G   F   S   V   N   M   R   S   L   D   T   L   L   A   C   818

3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG 4020
819  P   K   I   D   E   A   L   F   N   S   T   S   V   E   L   T   K   H   M   G   838

4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgtgtaactgaataatagctgacaaataatcag A TCG 4089
839  K   S   F   F   Y   K   I   L   R                                              S   848

4090 AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT 4149
849  S   L   A   S   F   A   Q   V   F   I   D   I   T   H   N   S   K   F   N   S   868

4150 TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA 4209
869  C   C   N   I   Y   R   L   G   Y   S   M   C   M   R   A   Q   A   Y   L   K   888

4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtgagtacttattttaactaga 4274
889  R   M   K   D   I   F   I   P   Q   R   M   F   I   T   D                        903

4275 aaagtcattaactaaccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AAG TTG GCC 4339
904                          L   L   N   V   I   G   R   K   I   W   K   K   L   A   917

4340 GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc 4401
918  E   I   L   G   Y   T   S   R   R   F   L   S   S   A   E   V   K   W            935

4402 ggtctcgagacttcagcaatattgacacatcag G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA 4468
936                                     L   F   C   L   G   M   R   D   G   L   K    946

4469 CCC TCT TTC AAA TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT 4528
947  P   S   F   K   Y   H   P   C   F   E   Q   L   I   Y   Q   F   Q   S   L   T   966

4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA AGA ATA 4588
967  D   L   I   K   P   L   R   P   V   L   R   Q   V   L   F   L   H   R   R   I   986

4589 GCT GAT TAA cgtcatttttaatttattatatacatcctttattactggtgtctttaaacaatattattactaagtata 4665
987  A   D   *                                                                         989
```

Figure 15
Page 5

```
4666 gctgacccccaaagcaagcatactataggatttttagtaaagtaaaacttaatcttgttattagttttgattgacttgtct 4745
4746 ttatcctttatacttttaagaaagattgacagtggttgctgactactgcccacatgcccattaaacgggagtggttaaaca 4825
4826 ttaaaagtaatacatgaggctaatcttcctttcatttagaataaggaaagtggttttctataatgaataatgcccgcacta 4905
4906 atgcaaaaaagacgaagattatctttctaaacaaggggggattaagcatatccgaaggaaaagagagtaatatacccagtgtt 4985
4986 gttgaagaaagcaaggataatttggaacaagcttctgcagatgacaggcttaaattttggtgaccgaattttggtaaaagc 5065
5066 cccaggttatccatggtggccggccttgctactgagacgaaaagaaactaaggatagtttgaatactaatagcttcattta 5145
5146 atgtcttatataaggttttgtttttttcctgacttcaacttttgcatgggtgaaaagaaatagtgttaagccattattggat 5225
5226 tccgaaatagccaaatttcttggttcctcaaagcggaagtcttaaagaacttattgaagcttatgaggcttcaaaaacttcc 5305
5306 tcctgatttaaaggaggaatccttccaccgatgaggaaatggatagcttatcagctgctgaggagaagcctaatttttgc 5385
5386 aaaaaagaaaatatcattgggagacatcttttgatgaatcagatgcggagagtatcttccagcggatccttgatgtcaata 5465
5466 acttctatttctgaaatgtatggtcctactgtcgcttcgacttttcgtagctttacgcagttaagtgaccaaaggtacc  5544
```

FIGURE 16

page 1

(Seq. ID. No. 1)

```
   1 gcagcgctgc gtcctgctgc gcacgtggga agccctggcc ccggccaccc ccgcgatgcc
  61 gcgcgctccc cgctgccgag ccgtgcgctc cctgctgcgc agccactacc gcgaggtgct
 121 gccgctggcc acgttcgtgc ggcgcctggg gccccagggc tggcggctgg tgcagcgcgg
 181 ggacccggcg gctttccgcg cgctggtggc ccagtgcctg gtgtgcgtgc cctgggacgc
 241 acggccgccc cccgccgccc cctccttccg ccaggtgtcc tgcctgaagg agctggtggc
 301 ccgagtgctg cagaggctgt gcgagcgcgg cgcgaagaac gtgctggcct tcggcttcgc
 361 gctgctggac ggggcccgcg ggggcccccc cgaggccttc accaccagcg tgcgcagcta
 421 cctgcccaac acggtgaccg acgcactgcg ggggagcggg gcgtgggggc tgctgctgcg
 481 ccgcgtgggc gacgacgtgc tggttcacct gctggcacgc tgcgcgctct ttgtgctggt
 541 ggctcccagc tgcgcctacc aggtgtgcgg gccgccgctg taccagctcg gcgctgccac
 601 tcaggcccgg cccccgccac acgctagtgg accccgaagg cgtctgggat gcgaacgggc
 661 ctggaaccat agcgtcaggg aggccggggt ccccctgggc ctgccagccc cgggtgcgag
 721 gaggcgcggg ggcagtgcca gccgaagtct gccgttgccc aagaggccca ggcgtggcgc
 781 tgcccctgag ccggagcgga cgcccgttgg gcaggggtcc tgggcccacc cgggcaggac
 841 gcgtggaccg agtgaccgtg gtttctgtgt ggtgtcacct gccagacccg ccgaagaagc
 901 cacctctttg gagggtgcgc tctctggcac gcgccactcc cacccatccg tgggccgcca
 961 gcaccacgcg ggcccccccat ccacatcgcg gccaccacgt ccctgggaca cgccttgtcc
1021 cccggtgtac gccgagacca agcacttcct ctactcctca ggcgacaagg agcagctgcg
1081 gccctccttc ctactcagct ctctgaggcc cagcctgact ggcgctcgga ggctcgtgga
1141 gaccatcttt ctgggttcca ggccctggat gccagggact ccccgcaggt tgccccgcct
1201 gccccagcgc tactggcaaa tgcggccccct gtttctggag ctgcttggga accacgcgca
1261 gtgcccctac ggggtgctcc tcaagacgca ctgcccgctg cgagctgcgg tcaccccagc
1321 agccggtgtc tgtgcccggg agaagcccca gggctctgtg gcggcccccg aggaggagga
1381 cacagacccc cgtcgcctgg tgcagctgct ccgccagcac agcagccccct ggcaggtgta
1441 cggcttcgtg cgggcctgcc tgcgccggct ggtgccccca ggcctctggg gctccaggca
1501 caacgaacgc cgcttcctca ggaacaccaa gaagttcatc tccctgggga agcatgccaa
1561 gctctcgctg caggagctga cgtggaagat gagcgtgcgg gactgcgctt ggctgcgcag
1621 gagcccaggg gttggctgtg ttccggccgc agagcaccgt ctgcgtgagg agatcctggc
1681 caagttcctg cactggctga tgagtgtgta cgtcgtcgag ctgctcaggt ctttctttta
1741 tgtcacggag accacgtttc aaaagaacag gctctttttc taccggaaga gtgtctggag
1801 caagttgcaa agcattggaa tcagacagca cttgaagagg gtgcagctgc gggagctgtc
1861 ggaagcagag gtcaggcagc atcgggaagc caggcccgcc ctgctgacgt ccagactccg
1921 cttcatcccc aagcctgacg ggctgcggcc gattgtgaac atggactacg tcgtgggagc
1981 cagaacgttc cgcagagaaa agagggccga gcgtctcacc tcgagggtga aggcactgtt
2041 cagcgtgctc aactacgagc gggcgcggcg ccccggcctc ctgggcgcct ctgtgctggg
2101 cctggacgat atccacaggg cctggcgcac cttcgtgctg cgtgtgcggg cccaggaccc
2161 gccgcctgag ctgtactttg tcaaggtgga tgtgacgggc gcgtacgaca ccatccccca
2221 ggacaggctc acggaggtca tcgccagcat catcaaaccc cagaacacgt actgcgtgcg
2281 tcggtatgcc gtggtccaga aggccgccca tgggcacgtc cgcaaggcct tcaagagcca
2341 cgtctctacc ttgacagacc tccagccgta catgcgacag ttcgtggctc acctgcagga
2401 gaccagcccg ctgagggatg ccgtcgtcat cgagcagagc tcctccctga atgaggccag
2461 cagtggcctc ttcgacgtct tcctacgctt catgtgccac cacgccgtgc gcatcagggg
2521 caagtcctac gtccagtgcc aggggatccc gcagggctcc atcctctcca cgctgctctg
2581 cagcctgtgc tacggcgaca tggagaacaa gctgtttgcg gggattcggc gggacgggct
2641 gctcctgcgt ttggtggatg atttcttgtt ggtgacacct cacctcaccc acgcgaaaac
2701 cttcctcagg accctggtcc gaggtgtccc tgagtatggc tgcgtggtga acttgcggaa
2761 gacagtggtg aacttccctg tagaagacga ggccctgggt ggcacggctt ttgttcagat
2821 gccggcccac ggcctattcc cctggtgcgg cctgctgctg gatacccgga ccctggaggt
2881 gcagagcgac tactccagct atgcccggac ctccatcaga gccagtctca ccttcaaccg
```

FIGURE 16
page 2
(Seq. ID. No. 1)

```
2941 cggcttcaag gctgggagga acatgcgtcg caaactcttt ggggtcttgc ggctgaagtg
3001 tcacagcctg tttctggatt tgcaggtgaa cagcctccag acggtgtgca ccaacatcta
3061 caagatcctc ctgctgcagg cgtacaggtt tcacgcatgt gtgctgcagc tcccatttca
3121 tcagcaagtt tggaagaacc ccacattttt cctgcgcgtc atctctgaca cggcctccct
3181 ctgctactcc atcctgaaag ccaagaacgc agggatgtcg ctgggggcca agggcgccgc
3241 cggcccctctg ccctccgagg ccgtgcagtg gctgtgccac caagcattcc tgctcaagct
3301 gactcgacac cgtgtcacct acgtgccact cctggggtca ctcaggacag cccagacgca
3361 gctgagtcgg aagctcccgg ggacgacgct gactgccctg gaggccgcag ccaacccggc
3421 actgccctca gacttcaaga ccatcctgga ctgatggcca cccgcccaca gccaggccga
3481 gagcagacac cagcagccct gtcacgccgg gctctacgtc ccagggaggg aggggcggcc
3541 cacacccagg cccgcaccgc tgggagtctg aggcctgagt gagtgtttgg ccgaggcctg
3601 catgtccggc tgaaggctga gtgtccggct gaggcctgag cgagtgtcca gccaagggct
3661 gagtgtccag cacacctgcc gtcttcactt ccccacaggc tggcgctcgg ctccacccca
3721 gggccagctt ttcctcacca ggagcccggc ttccactccc cacataggaa tagtccatcc
3781 ccagattcgc cattgttcac ccctcgccct gccctccttt gccttccacc cccaccatcc
3841 aggtggagac cctgagaagg accctgggag ctctgggaat ttggagtgac caaaggtgtg
3901 ccctgtacac aggcgaggac cctgcacctg gatgggggtc cctgtgggtc aaattgggggg
3961 gaggtgctgt gggagtaaaa tactgaatat atgagttttt cagttttgaa aaaaa
```

FIGURE 17
HUMAN TRT PROTEIN SEQUENCE
(SEQ. NO. 2)

MPRAPRCRAVRSLLRSHYREVLPLATFVRRLGPQGWRLVQRGDP

AAFRALVAQCLVCVPWDARPPPAAPSFRQVSCLKELVARVLQRLCERGAKNVLAFGFA

LLDGARGGPPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDDVLVHLLARCALFV

LVAPSCAYQVCGPPLYQLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPA

PGARRRGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRGPSDRGFCVVSPA

RPAEEATSLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPWDTPCPPVYAETKHFLYS

SGDKEQLRPSFLLSSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPL

FLELLGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPQGSVAAPEEEDTDPRRLVQ

LLRQHSSPWQVYGFVRACLRRLVPPGLWGSRHNERRFLRNTKKFISLGKHAKLSLQEL

TWKMSVRDCAWLRRSPGVGCVPAAEHRLREEILAKFLHWLMSVYVVELLRSFFYVTET

TFQKNRLFFYRKSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREARPALLTSRLRFI

PKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKALFSVLNYERARRPGLLGASVLG

LDDIHRAWRTFVLRVRAQDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYC

VRRYAVVQKAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVIEQSSSL

NEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPQGSILSTLLCSLCYGDMENKLFAG

IRRDGLLLRLVDDFLLVTPHLTHAKTFLRTLVRGVPEYGCVVNLRKTVVNFPVEDEAL

GGTAFVQMPAHGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGRNMRR

KLFGVLRLKCHSLFLDLQVNSLQTVCTNIYKILLLQAYRFHACVLQLPFHQQVWKNPT

FFLRVISDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVT

YVPLLGSLRTAQTQLSRKLPGTTLTALEAAANPALPSDFKTILD

FIGURE 18
Clone 712562
(SEQ ID NO. 3)

GGCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTT
TTATGTCACGGAGACCACGTTTCAAAAGAACAGGCTCTTTTTCTACCGGAAGAGTGTCTG
GAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCT
GTCGGAAGCAGAGGTCAGGCAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACT
CCGCTTCATCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGG
AGCCAGAACGTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACT
GTTCAGCGTGCTCAACTACGAGCGGGCGCGGCGCCCCGGCCTCCTGGGCGCCTCTGTGCT
GGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCCAGGA
CCCGCCGCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC
CCAGGACAGGCTCACGGAGGTCATCGCCAGCATCATCAAACCCCAGAACACGTACTGCGT
GCGTCGGTATGCCGTGGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAG
CCACGTCCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCT
GCAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGATTCGGCGGGACGGGC
TGCTCCTGCGTTTGGTGGATGATTTCTTGTTGGTGACACCTCACCTCACCCACGCGAAAA
CCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACTTGCGGA
AGACAGTGGTGAACTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGA
TGCCGGCCCACGGCCTATTCCCCTGGTGCGGCCTGCTGCTGGATACCCGGACCCTGGAGG
TGCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACC
GCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGT
GTCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT
ACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGCAGCTCCCATTTC
ATCAGCAAGTTTGGAAGAACCCCACATTTTTCCTGCGCGTCATCTCTGACACGGCCTCCC
TCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGATGTCGCTGGGGGCCAAGGGCGCCG
CCGGCC7TCTGCCCTCCGAGGCCGTGCAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGC
TGACTCGACACCGTGTCACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGC
AGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGCAGCCAACCCGG
CACTGCCCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCCACAGCCAGGCCG
AGAGCAGACACCAGCAGCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGC
CCACACCCAGGCCTGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCT
GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGC
TGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCC
AGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATC
CCCAGATTCGCCATTGTTCACCCCTCGCCCTGCCCTCCTTTGCCTTCCACCCCCACCATC
CAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGT
GCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGG
GGAGGTGCTGTGGGAGTAAAATACTGAATATATGAGTTTTTCAGTTTTG0AAAAAAAAAA
AAAAAAAAAAAAAAAA

FIGURE 19

SEQ ID NO. 10

MetSerValTyrValValGluLeuLeuArgSerPhePhe
TyrValThrGluThrThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrp
SerLysLeuGlnSerIleGlyIleArgGlnHisLeuLysArgValGlnLeuArgGluLeu
SerGluAlaGluValArgGlnHisArgGluAlaArgProAlaLeuLeuThrSerArgLeu
ArgPheIleProLysProAspGlyLeuArgProIleValAsnMetAspTyrValValGly
AlaArgThrPheArgArgGluLysArgAlaGluArgLeuThrSerArgValLysAlaLeu
PheSerValLeuAsnTyrGluArgAlaArgArgProGlyLeuLeuGlyAlaSerValLeu
GlyLeuAspAspIleHisArgAlaTrpArgThrPheValLeuArgValArgAlaGlnAsp
ProProProGluLeuTyrPheValLysValAspValThrGlyAlaTyrAspThrIlePro
GlnAspArgLeuThrGluValIleAlaSerIleIleLysProGlnAsnThrTyrCysVal
ArgArgTyrAlaValValGlnLysAlaAlaHisGlyHisValArgLysAlaPheLysSer
HisValLeuArgProValProGlyAspProAlaGlyLeuHisProLeuHisAlaAlaLeu
GlnProValLeuArgArgHisGlyGluGlnAlaValCysGlyAspSerAlaGlyArgAla
AlaProAlaPheGlyGly

FIGURE 20
Page 1

SEQUENCE NO. 4 (DNA) AND SEQUENCE NO. 5 (PROTEIN)
(TRANSLATION OF A Δ182 hTRT VARIANT)

```
                                                                      1
                                                                    met
GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCG ATG

                                        10
pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

              20                                                30
his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

                                        40
gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

              50                                                60
phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

                                        70
ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

              80                                                90
leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

                                       100
gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG

             110                                               120
ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC
```

FIGURE 20
Page 2

```
                                        130
tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

            140                                         150
trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

                                    160
leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

            170                                         180
ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

                                    190
thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

            200                                         210
leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

                                    220
val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

            230                                         240
ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

                                    250
ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

            260                                         270
ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT
```

FIGURE 20
Page 3

```
                                                    280
val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

            290                                                 300
gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

                                    310
gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

            320                                                 330
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

                                    340
leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

            350                                                 360
leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

                                    370
glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

            380                                                 390
arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

                                    400
leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

            410                                                 420
val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

                                    430
```

FIGURE 20
Page 4

```
ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

                440                                     450
ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

                                460
leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

                470                                     480
ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

                                490
his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

                500                                     510
leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

                                520
met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

                530                                     540
gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

                                550
ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

                560                                     570
leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

                                580
arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC
```

FIGURE 20
Page 5

```
                590                                                 600
ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

                                    610
ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

                620                                                 630
leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

                                    640
pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

                650                                                 660
arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

                                    670
phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG

                680                                                 690
gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

                                    700
thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG

                710                                                 720
tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

                                    730
gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG
```

FIGURE 20
Page 6

```
                740                                                 750
asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

                                    760
his gly his val arg lys ala phe lys ser his val leu arg pro
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA

                770                                                 780
val pro gly asp pro ala gly leu his pro leu his ala ala leu
GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG

                                    790
gln pro val leu arg arg his gly glu gln ala val cys gly asp
CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT

                800                                 807
ser ala gly arg ala ala pro ala phe gly gly OP
TCG GCG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTTGGT

GACACCTCACCTCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA

GTATGGCTGCGTGGTGAACTTGCGGAAGACAGTGGTGAACTTCCCTGTAGAAGACGAGGC

CCTGGGTGGCACGGCTTTTGTTCAGATGCCGGCCCACGGCCTATTCCCCTGGTGCGGCCT

GCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCGGACCTC

CATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAA

ACTCTTTGGGGTCTTGCGGCTGAAGTGTCACAGCCTGTTTCTGGATTTGCAGGTGAACAG

CCTCCAGACGGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCA

CGCATGTGTGCTGCAGCTCCCATTTCATCAGCAAGTTTGGAAGAACCCCACATTTTTCCT

GCGCGTCATCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG

GATGTCGCTGGGGGCCAAGGGCGCCGCCGGCCCTCTGCCCTCCGAGGCCGTGCAGTGGCT

GTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTCACCTACGTGCCACTCCT
```

FIGURE 20
Page 7

GGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGAC

TGCCCTGGAGGCCGCAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGACTG

ATGGCCACCCGCCCACAGCCAGGCCGAGAGCAGACACCAGCAGCCCTGTCACGCCGGGCT

CTACGTCCCAGGGAGGGAGGGGCGGCCCACACCCAGGCCCGCACCGCTGGGAGTCTGAGG

CCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG

GCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCC

CACAGGCTGGCGCTCGGCTCCACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTC

CACTCCCCACATAGGAATAGTCCATCCCCAGATTCGCCATTGTTCACCCCTCGCCCTGCC

CTCCTTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTC

TGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGAT

GGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATG

AGTTTTTCAGTTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 21

Genomic DNA insert of pGRN144

Seq. ID. No. 6

```
   1 CCATGGGACCCACTGCAGGGGCAGCTGGGAGGCTGCAGGCTTCAGGTCCCAGTGGGGTTG
     GGTACCCTGGGTGACGTCCCCGTCGACCCTCCGACGTCCGAAGTCCAGGGTCACCCCAAC

  61 CCATCTGCCAGTAGAAACCTGATGTAGAATCAGGGCGCGAGTGTGGACACTGTCCTGAAT
     GGTAGACGGTCATCTTTGGACTACATCTTAGTCCCGCGCTCACACCTGTGACAGGACTTA

 121 CTCAATGTCTCAGTGTGTGCTGAAACATGTAGAAATTAAAGTCCATCCCTCCTACTCTAC
     GAGTTACAGAGTCACACACGACTTTGTACATCTTTAATTTCAGGTAGGGAGGATGAGATG

 181 TGGGATTGAGCCCCTTCCCTATCCCCCCCCAGGGGCAGAGGAGTTCCTCTCACTCCTGTG
     ACCCTAACTCGGGGAAGGGATAGGGGGGGGTCCCCGTCTCCTCAAGGAGAGTGAGGACAC

 241 GAGGAAGGAATGATACTTTGTTATTTTTCACTGCTGGTACTGAATCCACTGTTTCATTTG
     CTCCTTCCTTACTATGAAACAATAAAAAGTGACGACCATGACTTAGGTGACAAAGTAAAC

                                 ****************************************
 301 TTGGTTTGTTTGTTTTGTTTTGAGAGGCGGTTTCACTCTTGTTGCTCAGGCTGGAGGGAG
     AACCAAACAAACAAAACAAAACTCTCCGCCAAAGTGAGAACAACGAGTCCGACCTCCCTC

     ************************************************************
 361 TGCAATGGCGCGATCTTGGCTTACTGCAGCCTCTGCCTCCCAGGTTCAAGTGATTCTCCT
     ACGTTACCGCGCTAGAACCGAATGACGTCGGAGACGGAGGGTCCAAGTTCACTAAGAGGA

                                                          alu
     ************************************************************
 421 GCTTCCGCCTCCCATTTGGCTGGGATTACAGGCACCCGCCACCATGCCCAGCTAATTTTT
     CGAAGGCGGAGGGTAAACCGACCCTAATGTCCGTGGGCGGTGGTACGGGTCGATTAAAAA

                                                            ==
     ************************************************************
 481 TGTATTTTTAGTAGAGACGGGGGTGGGGGTGGGGTTCACCATGTTGGCCAGGCTGGTCTC
     ACATAAAAATCATCTCTGCCCCCACCCCCACCCCAAGTGGTACAACCGGTCCGACCAGAG

          CAP
     ============>

     ************************************************************
 541 GAACTTCTGACCTCAGATGATCCACCTGCCTCTGCCTCCTAAAGTGCTGGGATTACAGGT
1      CTTGAAGACTGGAGTCTACTAGGTGGACGGAGACGGAGGATTTCACGACCCTAATGTCCA

     **********************
 601 GTGAGCCACCATGCCCAGCTCAGAATTTACTCTGTTTAGAAACATCTGGGTCTGAGGTAG
     CACTCGGTGGTACGGGTCGAGTCTTAAATGAGACAAATCTTTGTAGACCCAGACTCCATC
```

FIGURE 21 (continued, 2 of 6)

```
                                          CCAAT
                                  ***************>
 661 GAAGCTCACCCCACTCAAGTGTTGTGGTGTTTTAAGCCAATGATAGAATTTTTTTATTGT
     CTTCGAGTGGGGTGAGTTCACAACACCACAAAATTCGGTTACTATCTTAAAAAAATAACA

 721 TGTTAGAACACTCTTGATGTTTTACACTGTGATGACTAAGACATCATCAGCTTTTCAAAG
     ACAATCTTGTGAGAACTACAAAATGTGACACTACTGATTCTGTAGTAGTCGAAAAGTTTC

                                        CAP
                                  ***************>
 781 ACACACTAACTGCACCCATAATACTGGGGTGTCTTCTGGGTATCAGCGATCTTCATTGAA
     TGTGTGATTGACGTGGGTATTATGACCCCACAGAAGACCCATAGTCGCTAGAAGTAACTT

                                                          CAP
                                                  ************
 841 TGCCGGGAGGCGTTTCCTCGCCATGCACATGGTGTTAATTACTCCAGCATAATCTTCTGC
     ACGGCCCTCCGCAAAGGAGCGGTACGTGTACCACAATTAATGAGGTCGTATTAGAAGACG

     ***>
 901 TTCCATTTCTTCTCTTCCCTCTTTTAAAATTGTGTTTTCTATGTTGGCTTCTCTGCAGAG
     AAGGTAAAGAAGAGAAGGGAGAAAATTTTAACACAAAAGATACAACCGAAGAGACGTCTC

                                    CAP
                             ***************>
 961 AACCAGTGTAAGCTACAACTTAACTTTTGTTGGAACAAATTTTCCAAACCGCCCCTTTGC
     TTGGTCACATTCGATGTTGAATTGAAAACAACCTTGTTTAAAAGGTTTGGCGGGGAAACG

1021 CCTAGTGGCAGAGACAATTCACAAACACAGCCCTTTAAAAAGGCTTAGGGATCACTAAGG
     GGATCACCGTCTCTGTTAAGTGTTTGTGTCGGGAAATTTTTCCGAATCCCTAGTGATTCC

1081 GGATTTCTAGAAGAGCGACCCGTAATCCTTAAGTATTTACAAGACGAGGCTAACCTCCAG
     CCTAAAGATCTTCTCGCTGGGCATTAGGAATTCATAAATGTTCTGCTCCGATTGGAGGTC

1141 CGAGCGTGACAGCCCAGGGAGGGTGCGAGGCCTGTTCAAATGCTAAGCTTCCATAAATAA
     GCTCGCACTGTCGGGTCCCTCCCACGCTCCGGACAAGTTTACGATTCGAAGGTATTTATT

1201 AGCAAATTTCCTCCGGCAGTTTCTGGAAAGTAGGAAAGGTTAACATTTAAGGTTGCGTTT
     TCGTTTAAAGGAGGCCGTCAAAGACCTTTCATCCTTTCCAATTGTAAATTCCAACGCAAA

1261 GTTAGCATTTCAGTGTTTGCCGACCTCAGCTAACAGCATCCCTGCAAGGCCTCGGGAGAC
     CAATCGTAAAGTCACAAACGGCTGGAGTCGATTGTCGTAGGGACGTTCCGGAGCCCTCTG

1321 CCAGAAGTTTCTCGCCCCTTAGATCCAAACTTGAGCAACCCGGAGTCTGGATTCCTGGGA
     GGTCTTCAAAGAGCGGGGAATCTAGGTTTGAACTCGTTGGGCCTCAGACCTAAGGACCCT

                      TopoII
                ****************>
1381 AGTCCTCAGCTGTCCTGCGGTTGTGCCGGGGCCCCAGGTCTGGAGGGGACCAGTGGCCGT
     TCAGGAGTCGACAGGACGCCAACACGGCCCCGGGGTCCAGACCTCCCCTGGTCACCGGCA

1441 GTGGCTTCTACTGCTGGGCTGGAAGTCGGGCCTCCTAGCTCTGCAGTCCGAGGCTTGGAG
     CACCGAAGATGACGACCCGACCTTCAGCCCGGAGGATCGAGACGTCAGGCTCCGAACCTC
```

FIGURE 21 (continued, 3 of 6)

```
1501 CCAGGTGCCTGGACCCCGAGGCTGCCCTCCACCCTGTGCGGGCGGGATGTGACCAGATGT
     GGTCCACGGACCTGGGGCTCCGACGGGAGGTGGGACACGCCCGCCCTACACTGGTCTACA

1561 TGGCCTCATCTGCCAGACAGAGTGCCGGGGCCCAGGGTCAAGGCCGTTGTGGCTGGTGTG
     ACCGGAGTAGACGGTCTGTCTCACGGCCCCGGGTCCCAGTTCCGGCAACACCGACCACAC

1621 AGGCGCCCGGTGCGCGGCCAGCAGGAGCGCCTGGCTCCATTTCCCACCCTTTCTCGACGG
     TCCGCGGGCCACGCGCCGGTCGTCCTCGCGGACCGAGGTAAAGGGTGGGAAAGAGCTGCC

1681 GACCGCCCCGGTGGGTGATTAACAGATATTGGGGTGGTTTGCTCATGGTGGGGACCCCTT
     CTGGCGGGGCCACCCACTAATTGTCTATAACCCCACCAAACGAGTACCACCCCTGGGGAA

1741 CGCCGCCTGAGAACCTGCAAAGAGAAATGACGGGCCTGTGTCAAGGAGCCCAAGTCGCGG
     GCGGCGGACTCTTGGACGTTTCTCTTTACTGCCCGGACACAGTTCCTCGGGTTCAGCGCC

1801 GGAAGTGTTGCAGGGAGGCACTCCGGGAGGTCCCGCGTGCCCGTCCAGGGAGCAATGCGT
     CCTTCACAACGTCCCTCCGTGAGGCCCTCCAGGGCGCACGGGCAGGTCCCTCGTTACGCA

1861 CCTCGGGTTCGTCCCCAGCCGCGTCTACGCGCCTCCGTCCTCCCCTTCACGTCCGGCATT
     GGAGCCCAAGCAGGGGTCGGCGCAGATGCGCGGAGGCAGGAGGGGAAGTGCAGGCCGTAA

1921 CGTGGTGCCCCGGAGCCCGACGCCCCGCGTCCGGACCTGGAGGCAGCCCTGGGTCTCCGGA
     GCACCACGGGCCTCGGGCTGCGGGGCGCAGGCCTGGACCTCCGTCGGGACCCAGAGGCCT

1981 TCAGGCCAGCGGCCAAAGGGTCGCCGCACGCACCTGTTCCCAGGGCCTCCACATCATGGC
     AGTCCGGTCGCCGGTTTCCCAGCGGCGTGCGTGGACAAGGGTCCCGGAGGTGTAGTACCG

2041 CCCTCCCTCGGGTTACCCCACAGCCTAGGCCGATTCGACCTCTCTCCGCTGGGGCCCTCG
     GGGAGGGAGCCCAATGGGGTGTCGGATCCGGCTAAGCTGGAGAGAGGCGACCCCGGGAGC

                                                  Sp1
                                                 ********
2101 CTGGCGTCCCTGCACCCTGGGAGCGCGAGCGGCGCGCGGGCGGGGAAGCGCGGCCCAGAC
     GACCGCAGGGACGTGGGACCCTCGCGCTCGCCGCGCGCCCGCCCCTTCGCGCCGGGTCTG

2161 CCCCGGGTCCGCCCGGAGCAGCTGCGCTGTCGGGGCCAGGCCGGGCTCCCAGTGGATTCG
     GGGGCCCAGGCGGGCCTCGTCGACGCGACAGCCCCGGTCCGGCCCGAGGGTCACCTAAGC

2221 CGGGCAACAGACGCCCAGGACCGCGCTTCCCACGTGGCGGAGGGACTGGGGACCCGGGCA
     GCCCGTTGTCTGCGGGTCCTGGCGCGAAGGGTGCACCGCCTCCCTGACCCCTGGGCCCGT

                                                        Sp1
                                                       ========
                                                  E2F
                                                 ********
2281 CCGGTCCTGCCCCTTCACCTTCCAGCTCCGCCTCGTCCGCGCGGAACCCCGCCCCGTCCC
     GGCCAGGACGGGGAAGTGGAAGGTCGAGGCGGAGCAGGCGCGCCTTGGGGCGGGGCAGGG

2341 GAACCCTTCCCGGGTCCCCGGCCCAGCCCCTTCCGGGCCATCCCAGCCCGTCCCGTTCCT
     CTTGGGAAGGGCCCAGGGGCCGGGTCGGGGAAGGCCCGGTAGGGTCGGGCAGGGCAAGGA
```

FIGURE 21 (continued, 4 of 6)

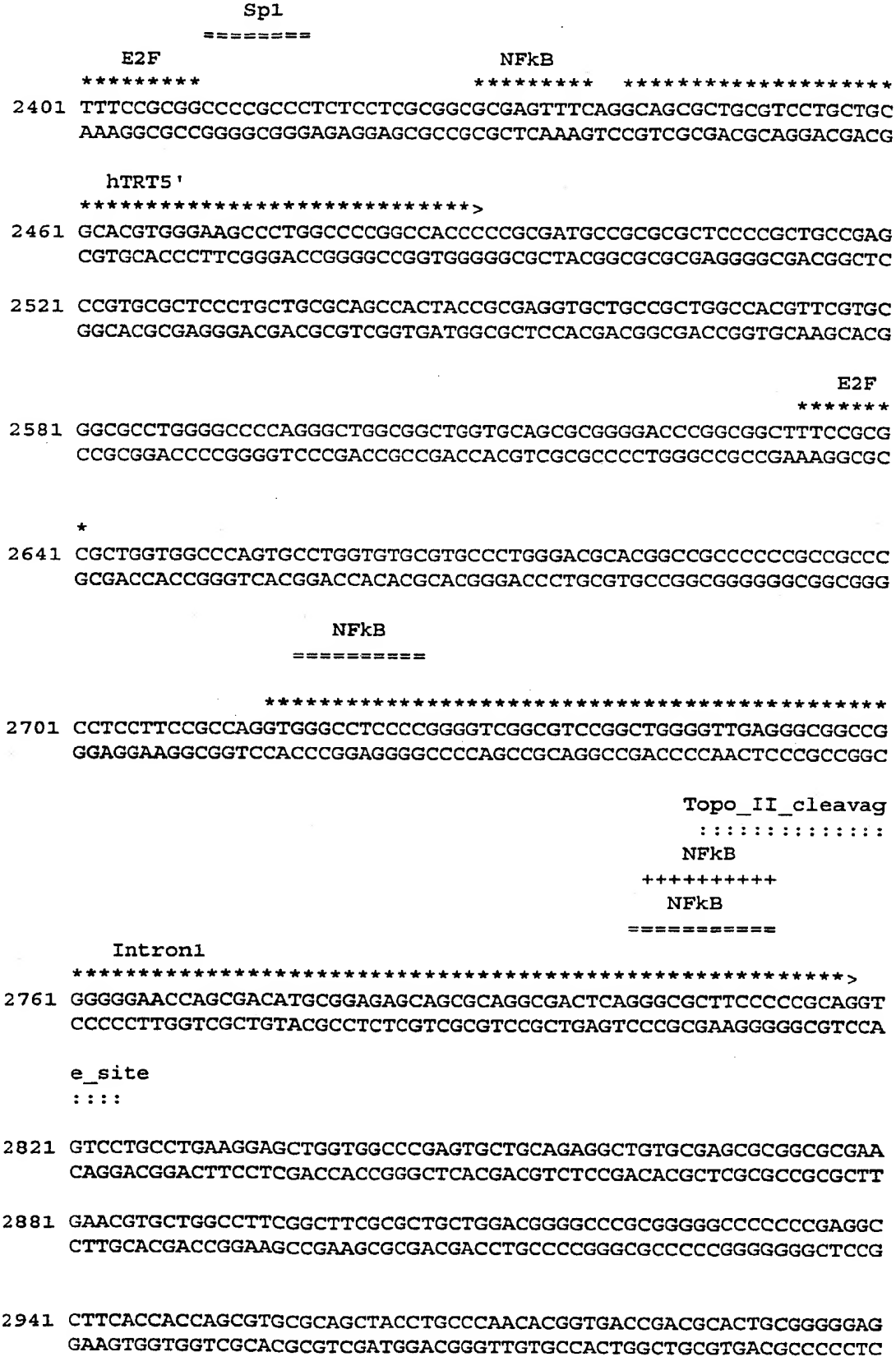


FIGURE 21 (continued, 5 of 6)

```
3001 CGGGGCGTGGGGGCTGCTGCTGCGCCGCGTGGGCGACGACGTGCTGGTTCACCTGCTGGC
     GCCCCGCACCCCCGACGACGACGCGGCGCACCCGCTGCTGCACGACCAAGTGGACGACCG

3061 ACGCTGCGCGCTCTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCC
     TGCGACGCGCGAGAAACACGACCACCGAGGGTCGACGCGGATGGTCCACACGCCCGGCGG

3121 GCTGTACCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTAGTGGACCCCG
     CGACATGGTCGAGCCGCGACGGTGAGTCCGGGCCGGGGGCGGTGTGCGATCACCTGGGGC

3181 AAGGCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCT
     TTCCGCAGACCCTACGCTTGCCCGGACCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGA

3241 GGGCCTGCCAGCCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTT
     CCCGGACGGTCGGGGCCCACGCTCCTCCGCGCCCCCGTCACGGTCGGCTTCAGACGGCAA

3301 GCCCAAGAGGCCCAGGCGTGGCGCTGCCCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGG
     CGGGTTCTCCGGGTCCGCACCGCGACGGGGACTCGGCCTCGCCTGCGGGCAACCCGTCCC

3361 GTCCTGGGCCCACCCGGGCAGGACGCGTGGACCGAGTGACCGTGGTTTCTGTGTGGTGTC
     CAGGACCCGGGTGGGCCCGTCCTGCGCACCTGGCTCACTGGCACCAAAGACACACCACAG

3421 ACCTGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCA
     TGGACGGTCTGGGCGGCTTCTTCGGTGGAGAAACCTCCCACGCGAGAGACCGTGCGCGGT

3481 CTCCCACCCATCCGTGGGCCGCCAGCACCACGCGGGCCCCCCATCCACATCGCGGCCACC
     GAGGGTGGGTAGGCACCCGGCGGTCGTGGTGCGCCCGGGGGGTAGGTGTAGCGCCGGTGG

3541 ACGTCCCTGGGACACGCCTTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTC
     TGCAGGGACCCTGTGCGGAACAGGGGGCCACATGCGGCTCTGGTTCGTGAAGGAGATGAG

3601 CTCAGGCGACAAGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCT
     GAGTCCGCTGTTCCTCGTCGACGCCGGGAGGAAGGATGAGTCGAGAGACTCCGGGTCGGA

3661 GACTGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCCAGGCCCTGGATGCCAGG
     CTGACCGCGAGCCTCCGAGCACCTCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCC

3721 GACTCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGCCCCTGTTTCT
     CTGAGGGGCGTCCAACGGGGCGGACGGGGTCGCGATGACCGTTTACGCCGGGGACAAAGA

3781 GGAGCTGCTTGGGAACCACGCGCAGTGCCCCTACGGGGTGCTCCTCAAGACGCACTGCCC
     CCTCGACGAACCCTTGGTGCGCGTCACGGGGATGCCCCACGAGGAGTTCTGCGTGACGGG

3841 GCTGCGAGCTGCGGTCACCCCAGCAGCCGGTGTCTGTGCCCGGGAGAAGCCCCAGGGCTC
     CGACGCTCGACGCCAGTGGGGTCGTCGGCCACAGACACGGGCCCTCTTCGGGGTCCCGAG

3901 TGTGGCGGCCCCCGAGGAGGAGGACACAGACCCCCGTCGCCTGGTGCAGCTGCTCCGCCA
     ACACCGCCGGGGGCTCCTCCTCCTGTGTCTGGGGGCAGCGGACCACGTCGACGAGGCGGT

3961 GCACAGCAGCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCTGCGCCGGCTGGTGCC
     CGTGTCGTCGGGGACCGTCCACATGCCGAAGCACGCCCGGACGGACGCGGCCGACCACGG

4021 CCCAGGCCTCTGGGGCTCCAGGCACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTT
     GGGTCCGGAGACCCCGAGGTCCGTGTTGCTTGCGGCGAAGGAGTCCTTGTGGTTCTTCAA
```

FIGURE 21 (continued, 6 of 6)

```
4081 CATCTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGT
     GTAGAGGGACCCCTTCGTACGGTTCGAGAGCGACGTCCTCGACTGCACCTTCTACTCGCA

                                          ****************************
4141 GCGGGACTGCGCTTGGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTCGAGGGCCC
     CGCCCTGACGCGAACCGACGCGTCCTCGGGTCCACTCCTCCACCACCGGCAGCTCCCGGG

                                                          Intron2
     ************************************************************
4201 AGGCCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCCCTGGTCCT
     TCCGGGGTCTCGACTTACGTCATCCCCGAGTCTTTTCCCCCGTCCGTCTCGGGACCAGGA

     ************************************************************
4261 CCTGTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTCGAGTGGACACG
     GGACAGAGGTAGCAGTGCACCCGTGTGCACCGAAAAGCGAGTCCTGCAGCTCACCTGTGC

     *****>
4321 GTGATCGAGGTCGAC
     CACTAGCTCCAGCTG
```

Figure 22

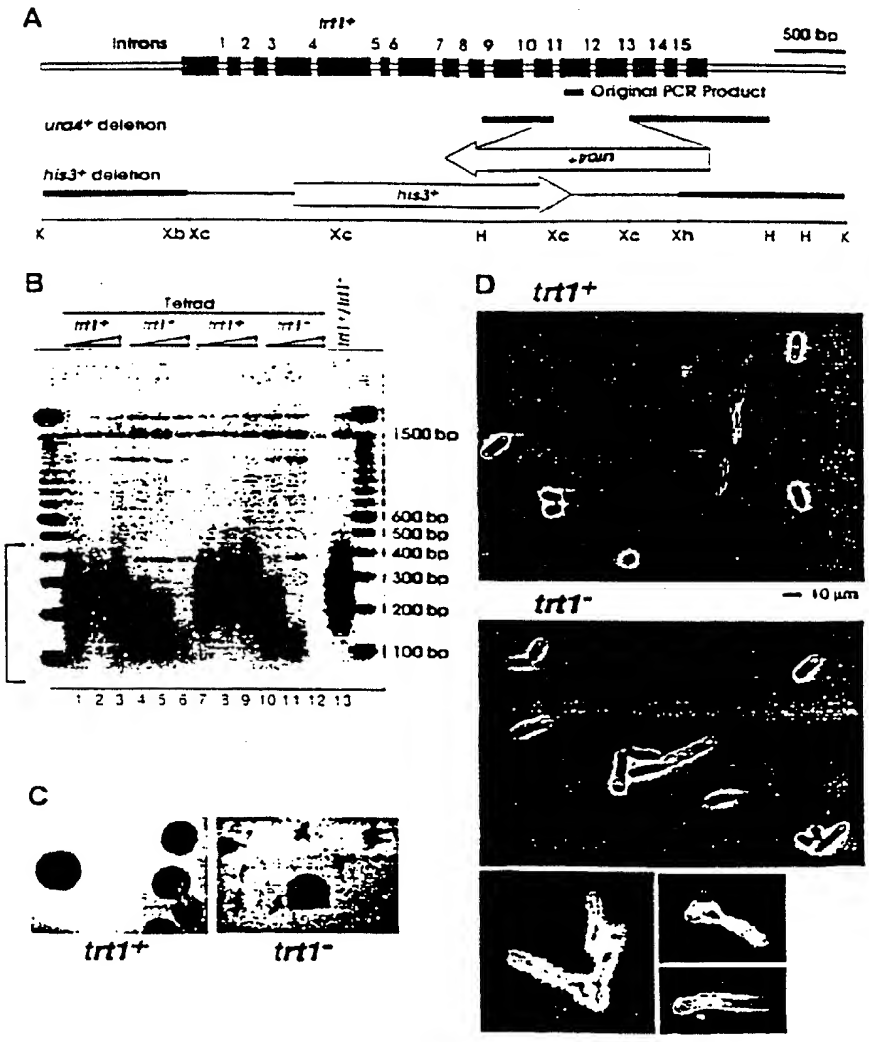

A
trt1+
Introns
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
500 bp
Original PCR Product
ura4+ deletion
+pUR
his3+ deletion
his3+
K Xb Xc Xc H Xc Xc Xh H H K
B
Tetrad
trt1+ trt1- trt1+ trt1-
trt1+/trt1-
1500 bp
600 bp
500 bp
400 bp
300 bp
200 bp
100 bp
1 2 3 4 5 6 7 8 9 10 11 12 13
C
trt1+
trt1-
D
trt1+
trt1-
10 µm

FIGURE 23
EST AA281296
(Seq. ID. No. 8)

gc
caagttcctg cactggctga tgagtgtgta cgtcgtcgag ctgctcaggt ctttctttta
tgtcacggag accacgtttc aaaagaacag gctctttttc taccggaaga gtgtctggag
caagttgcaa agcattggaa tcagacagca cttgaagagg gtgcagctgc gggacgtgtc
ggaagcagag gtcaggcagc atcgggaagc caggcccgcc ctgctgacgt ccagactccg
cttcatcccc aagcctgacg ggctgcggcc gattgtgaac atggactacg tcgtgggagc
cagaacgttc cgcagagaaa agagggccga gcgtctcacc tcgagggtga aggcactgtt
cagcgtgctc aactacgagc gggcgcg

FIGURE 24 -
(Seq. ID. No. 9)

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAG
ACCAGCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGC
CAGCAGTGGCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCAT
CAGGGGCAAGTC

Figure 25

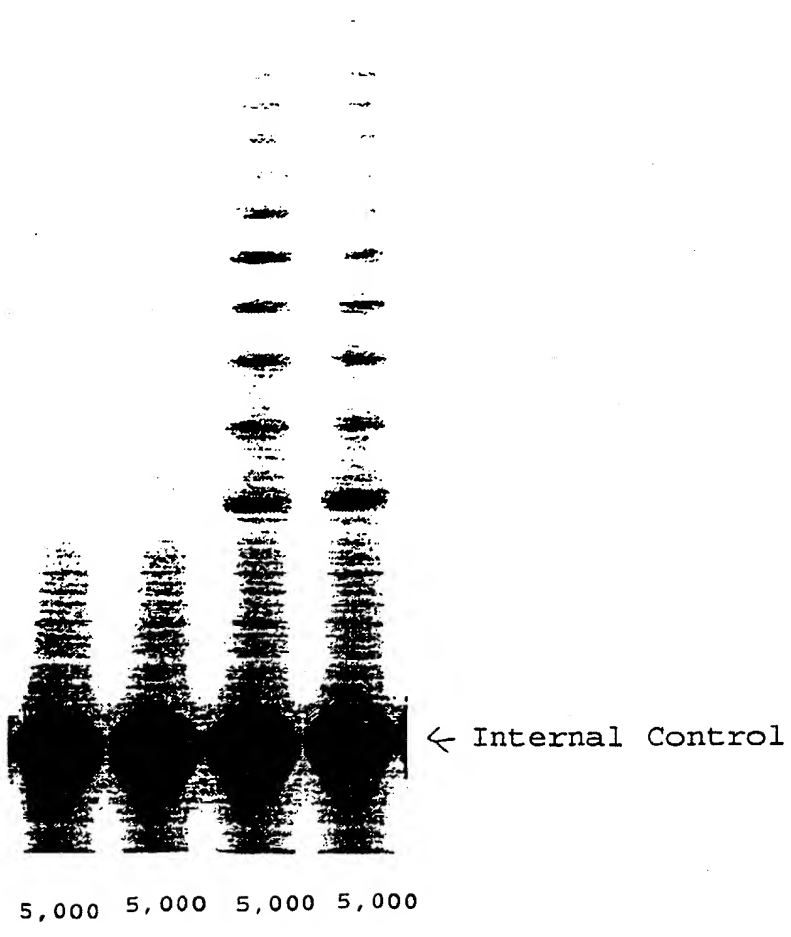

pBB5212
pGRN133
← Internal Control
Approximate Cell No.
5,000 5,000 5,000 5,000

Figure 26

PANEL A

PANEL B

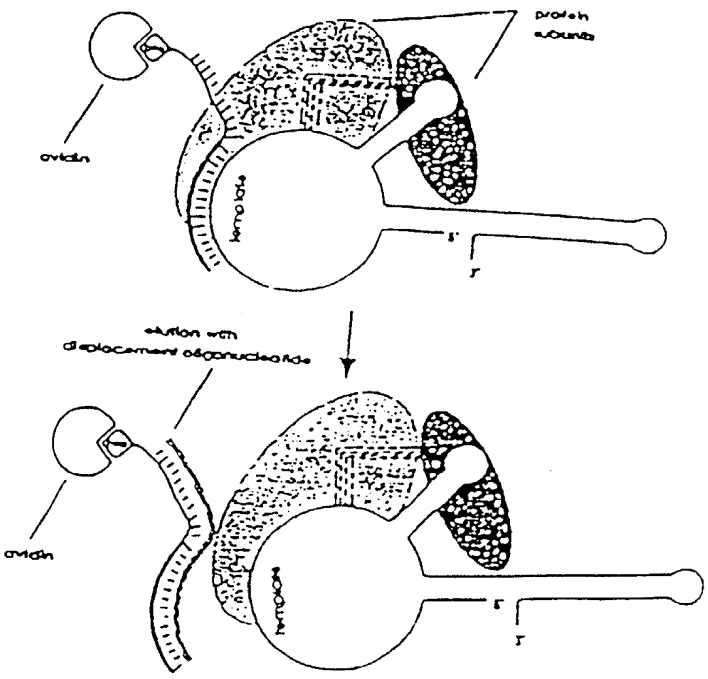


Figure 27

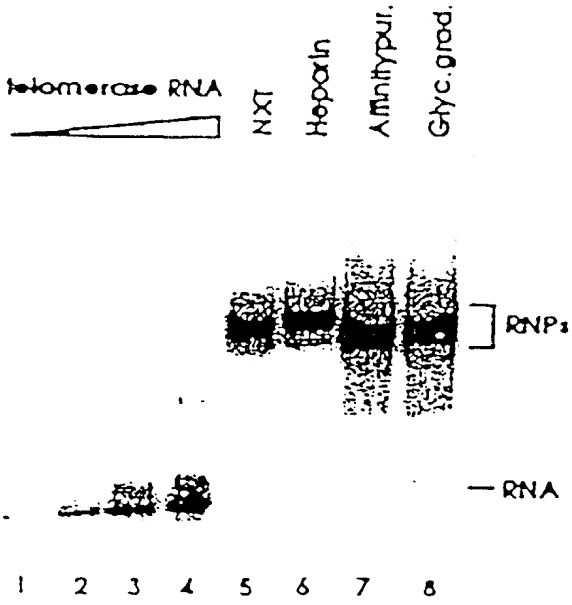

telomerase RNA
NXT
Heparin
Amithypur.
Glyc.grad.
RNP
RNA
1 2 3 4 5 6 7 8

Figure 28

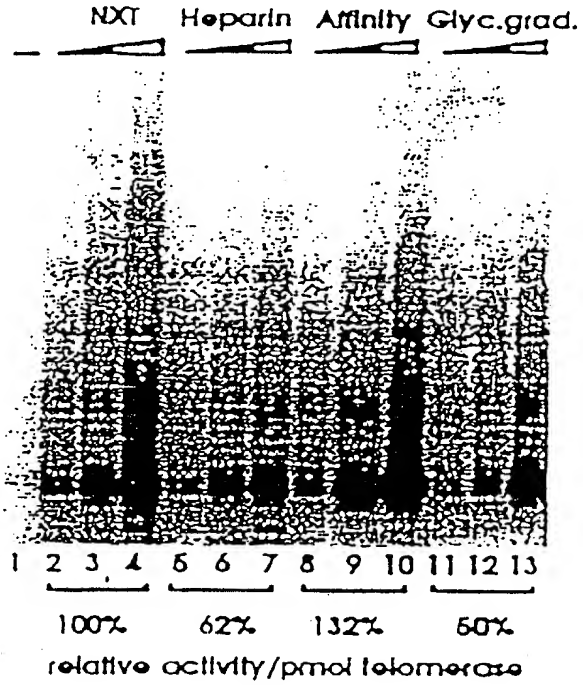

NXT
Heparin
Affinity
Glyc.grad.
1 2 3 4 5 6 7 8 9 10 11 12 13
100%
62%
132%
50%
relative activity/pmol telomerase

Figure 29

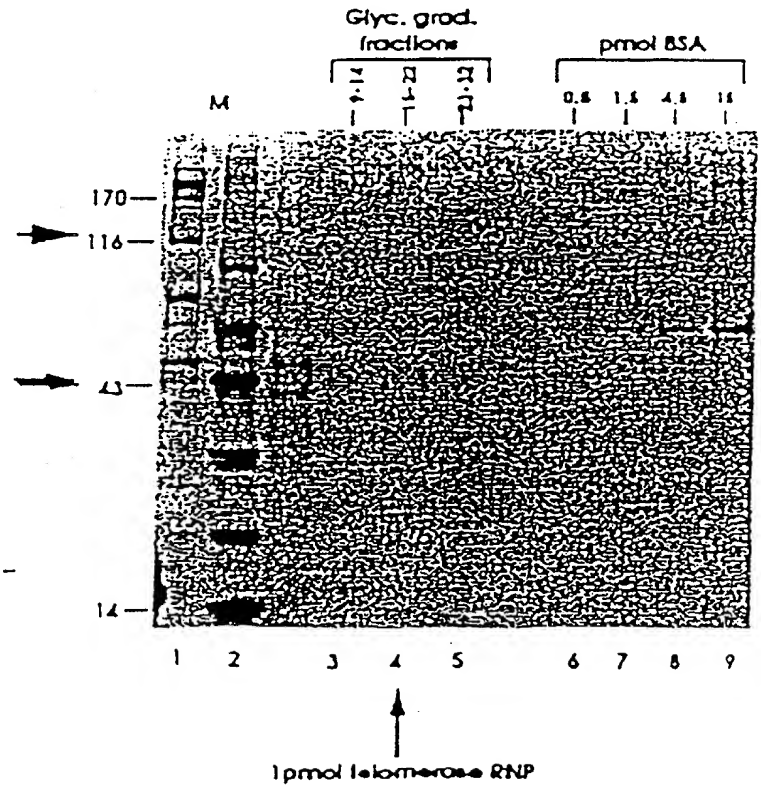

Glyc. grad. fractions
pmol BSA
M
0.5
1.5
4.5
15
170
116
43
14
1
2
3
4
5
6
7
8
9
1pmol telomerase RNP

Figure 30

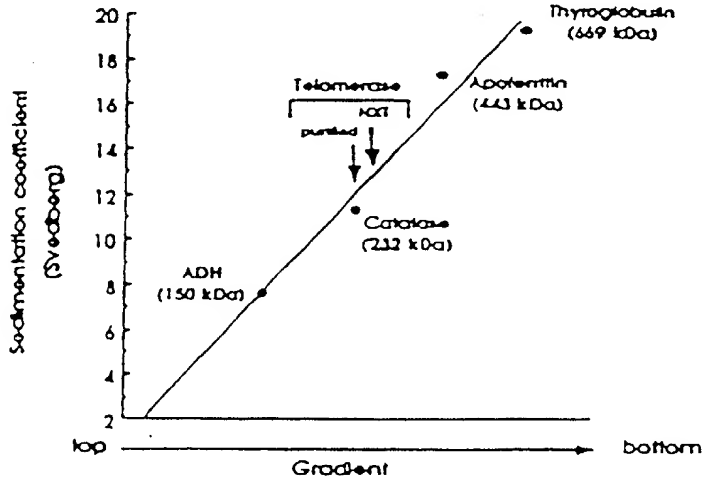

Sedimentation coefficient (Svedberg)
20
18
16
14
12
10
8
6
4
2
Thyroglobulin (669 kDa)
Telomerase
HeLa
purified
Apoferritin (443 kDa)
Catalase (232 kDa)
ADH (150 kDa)
top
bottom
Gradient

Figure 31

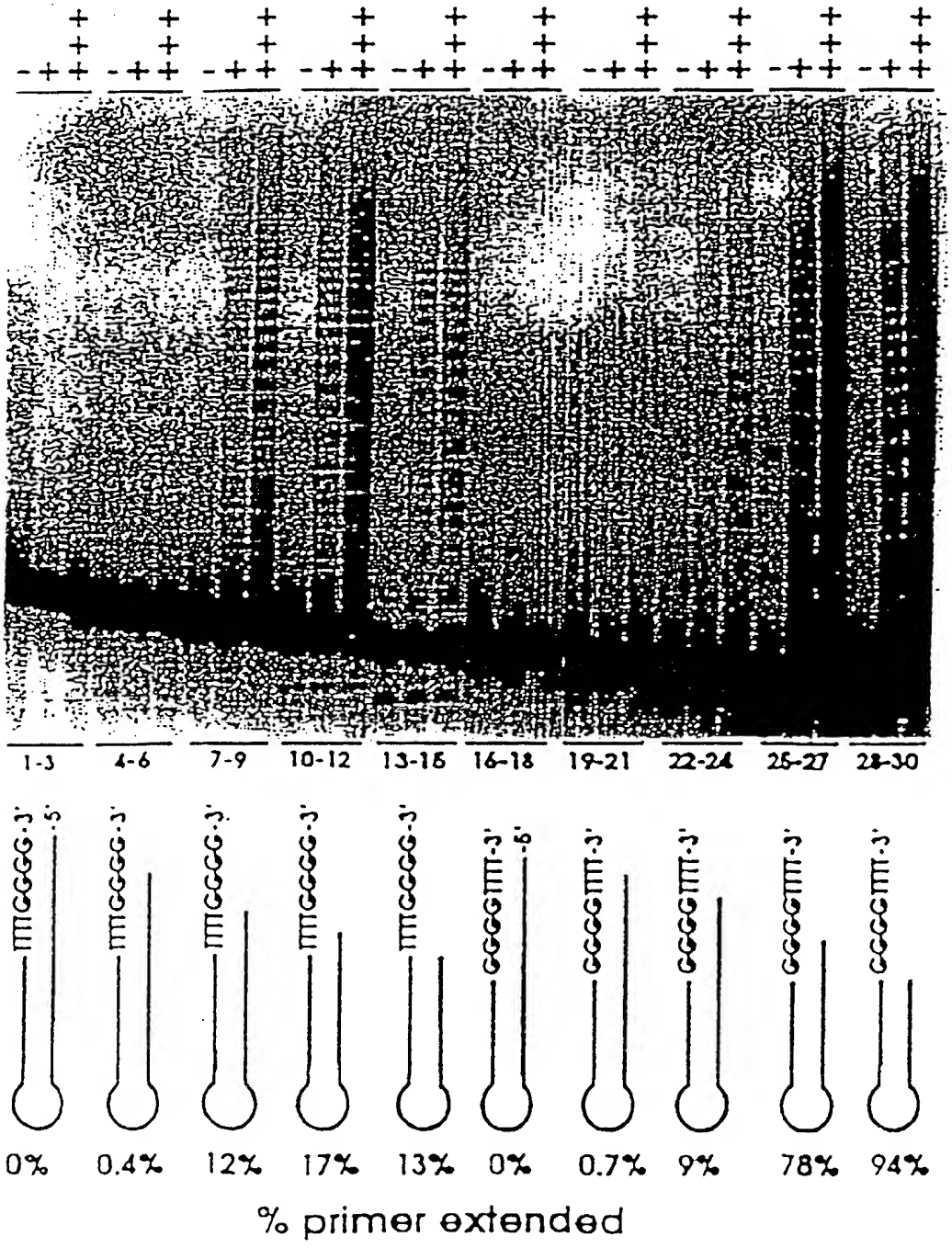

telomerase:
1-3
4-6
7-9
10-12
13-15
16-18
19-21
22-24
25-27
28-30
0%
0.4%
12%
17%
13%
0%
0.7%
9%
78%
94%
% primer extended

Figure 32

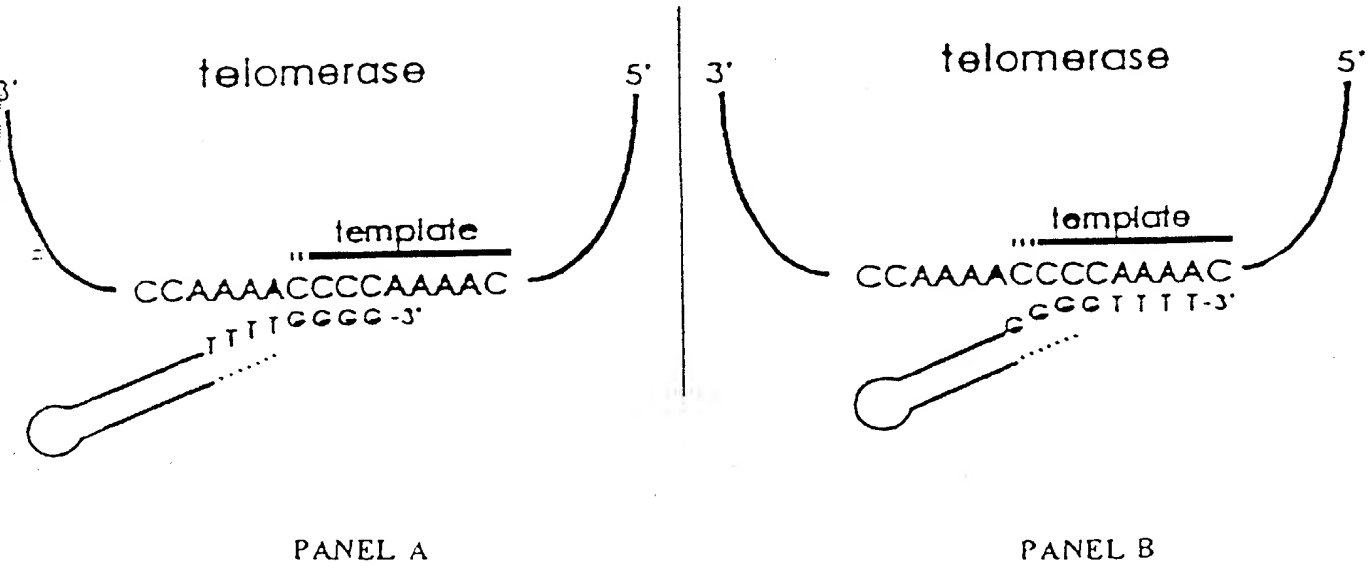

telomerase
3'
5'
template
CCAAAACCCCAAAAC
TTTTGGGG-3'
PANEL A
telomerase
3'
5'
template
CCAAAACCCCAAAAC
GGGTTTTT-3'
PANEL B

Figure 33

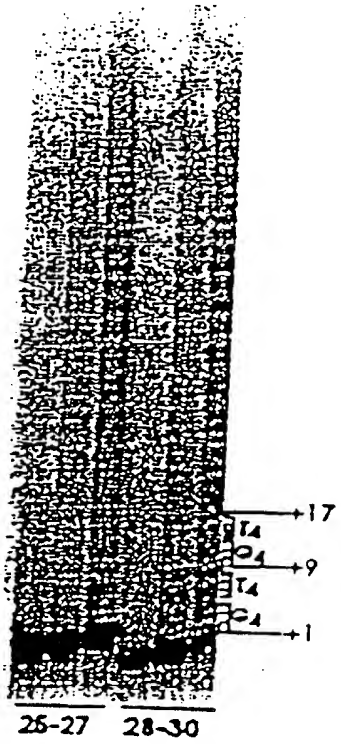


Figure 34

```
   1 CCCCAAAACC CCAAAACCCC AAAACCCCTA TAAAAAAAGA AAAAATTGAG
  51 GTAGTTTAGA AATAAAATAT TATTCCCGCA CAAATGGAGA TGGATATTGA
 101 TTTGGATGAT ATAGAAAATT TACTTCCTAA TACATTCAAC AAGTATAGCA
 151 GCTCTTGTAG TGACAAGAAA GGATGCAAAA CATTGAAATC TGGCTCGAAA
 201 TCGCCTTCAT TGACTATTCC AAAGTTGCAA AAACAATTAG AGTTCTACTT
 251 CTCGGATGCA AATCTTTATA ACGATTCTTT CTTGAGAAAA TTAGTTTTAA
 301 AAAGCGGAGA GCAAAGAGTA GAAATTGAAA CATTACTAAT GTTTAAATAA
 351 AATCAGGTAA TGAGGATTAT TCTATTTTTT AGATCACTTC TTAAGGAGCA
 401 TTATGGAGAA AATTACTTAA TACTAAAAGG TAAACAGTTT GGATTATTTC
 451 CCTAGCCAAC AATGATGAGT ATATTAAATT CATATGAGAA TGAGTCAAAG
 501 GATCTCGATA CATCAGACTT ACCAAAGACA AACTCGCTAT AAAACGCAAG
 551 AAAAAGTTTG ATAATCGAAC AGCAGAAGAA CTTATTGCAT TTACTATTCG
 601 TATGGGTTTT ATTACAATTG TTTTAGGTAT CGACGGTGAA CTCCCGAGTC
 651 TTGAGACAAT TGAAAAAGCT GTTTACAACT GAAGGAATCG CAGTTCTGAA
 701 AGTTCTGATG TGTATGCCAT TATTTTGTGA ATTAATCTCA AATATCTTAT
 751 CTCAATTTAA TGGATAGCTA TAGAAACAAA CCAAATAAAC CATGCAAGTT
 801 TAATGGAATA TACGTTAAAT CCTTTGGGAC AAATGCACAC TGAATTTATA
 851 TTGGATTCTT AAAGCATAGA TACACAGAAT GCTTTAGAGA CTGATTTAGC
 901 TTACAACAGA TTACCTGTTT TGATTACTCT TGCTCATCTC TTATATCTTT
 951 AAAAGAAGCA GGCGAAATGA AAAGAAGACT AAAGAAAGAG ATTTCAAAAT
1001 TTGTTGATTC TTCTGTAACC GGAATTAACA ACAAGAATAT TAGCAACGAA
1051 AAAGAAGAAG AGCTATCACA ATCCTGATTC TTAAAGATTT CAAAAATTCC
1101 AGGTAAGAGA GATACATTCA TTAAAATTCA TATATTATAG TTTTTCATTT
1151 CACAGCTGTT ATTTTCTTTT ATCTTAACAA TATTTTTTGA TTAGCTGGAA
1201 GTAAAAAGTA TCAAATAAGA GAAGCGCTAG ACTGAGGTAA CTTAGCTTAT
1251 TCACATTCAT AGATCGACCT TCATATATCC AATACGATGA TAAGGAAACA
1301 GCAGTCATCC GTTTTAAAAA TAGTGCTATG AGGACTAAAT TTTTAGAGTC
1351 AAGAAATGGA GCCGAAATCT TAATCAAAAA GAATTGCGTC GATATTGCAA
1401 AAGAATCGAA CTCTAAATCT TTCGTTAATA AGTATTACCA ATCTTGATTG
1451 ATTGAAGAGA TTGACGAGGC AACTGCACAG AAGATCATTA AAGAAATAAA
1501 GTAACTTTTA TTAATTAGAG AATAAACTAA ATTACTAATA TAGAGATCAG
1551 CGATCTTCAA TTGACGAAAT AAAAGCTGAA CTAAAGTTAG ACAATAAAAA
1601 ATACAAACCT TGGTCAAAAT ATTGAGGAAG GAAAAGAAGA CCAGTTAGCA
1651 AAAGAAAAAA TAAGGCAATA AATAAAATGA GTACAGAAGT GAAGAAATAA
1701 AAGATTTATT TTTTTCAATA ATTTATTGAA AAGAGGGGTT TTGGGGTTTT
1751 GGGGTTTTGG GG
```

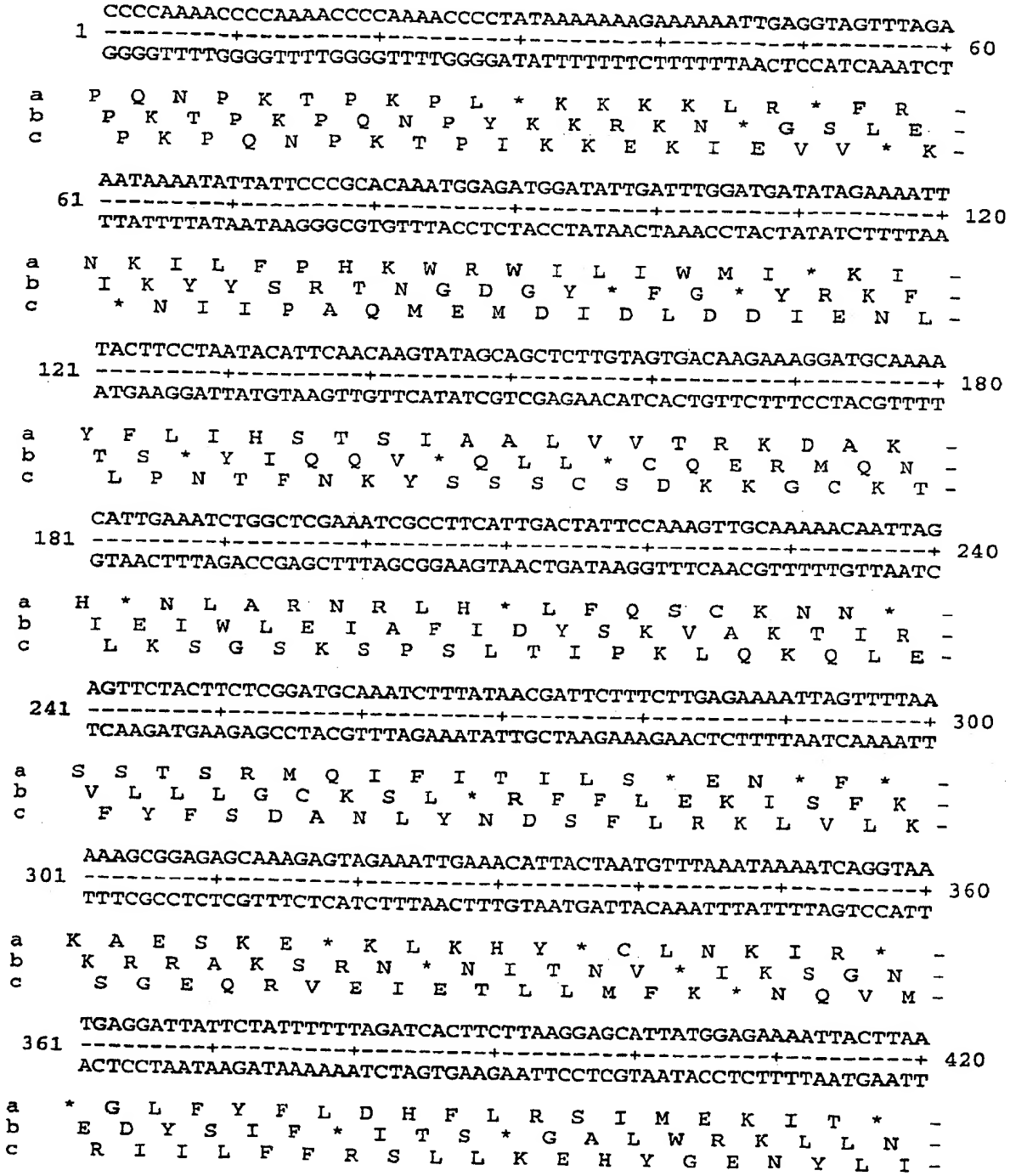

CCCCAAAACCCCAAAACCCCAAAACCCCTATAAAAAAAGAAAAAATTGAGGTAGTTTAGA
GGGGTTTTGGGGTTTTGGGGTTTTGGGGATATTTTTTTCTTTTTTAACTCCATCAAATCT
P Q N P K T P K P L * K K K K L R * F R
P K T P K P Q N P Y K K R K N * G S L E
P K P Q N P K T P I K K E K I E V V * K
AATAAAATATTATTCCCGCACAAATGGAGATGGATATTGATTTGGATGATATAGAAAATT
TTATTTTATAATAAGGGCGTGTTTACCTCTACCTATAACTAAACCTACTATATCTTTTAA
N K I L F P H K W R W I L I W M I * K I
I K Y Y S R T N G D G Y * F G * Y R K F
* N I I P A Q M E M D I D L D D I E N L
TACTTCCTAATACATTCAACAAGTATAGCAGCTCTTGTAGTGACAAGAAAGGATGCAAAA
ATGAAGGATTATGTAAGTTGTTCATATCGTCGAGAACATCACTGTTCTTTCCTACGTTTT
Y F L I H S T S I A A L V V T R K D A K
T S * Y I Q Q V * Q L L * C Q E R M Q N
L P N T F N K Y S S S C S D K K G C K T
CATTGAAATCTGGCTCGAAATCGCCTTCATTGACTATTCCAAAGTTGCAAAAACAATTAG
GTAACTTTAGACCGAGCTTTAGCGGAAGTAACTGATAAGGTTTCAACGTTTTTGTTAATC
H * N L A R N R L H * L F Q S C K N N *
I E I W L E I A F I D Y S K V A K T I R
L K S G S K S P S L T I P K L Q K Q L E
AGTTCTACTTCTCGGATGCAAATCTTTATAACGATTCTTTCTTGAGAAAATTAGTTTTAA
TCAAGATGAAGAGCCTACGTTTAGAAATATTGCTAAGAAAGAACTCTTTTAATCAAAATT
S S T S R M Q I F I T I L S * E N * F *
V L L L G C K S L * R F F L E K I S F K
F Y F S D A N L Y N D S F L R K L V L K
AAAGCGGAGAGCAAAGAGTAGAAATTGAAACATTACTAATGTTTAAATAAAATCAGGTAA
TTTCGCCTCTCGTTTCTCATCTTTAACTTTGTAATGATTACAAATTTATTTTAGTCCATT
K A E S K E * K L K H Y * C L N K I R *
K R R A K S R N * N I T N V * I K S G N
S G E Q R V E I E T L L M F K * N Q V M
TGAGGATTATTCTATTTTTTAGATCACTTCTTAAGGAGCATTATGGAGAAAATTACTTAA
ACTCCTAATAAGATAAAAAATCTAGTGAAGAATTCCTCGTAATACCTCTTTTAATGAATT
* G L F Y F L D H F L R S I M E K I T *
E D Y S I F * I T S * G A L W R K L L N
R I I L F F R S L L K E H Y G E N Y L I

FIG. 35

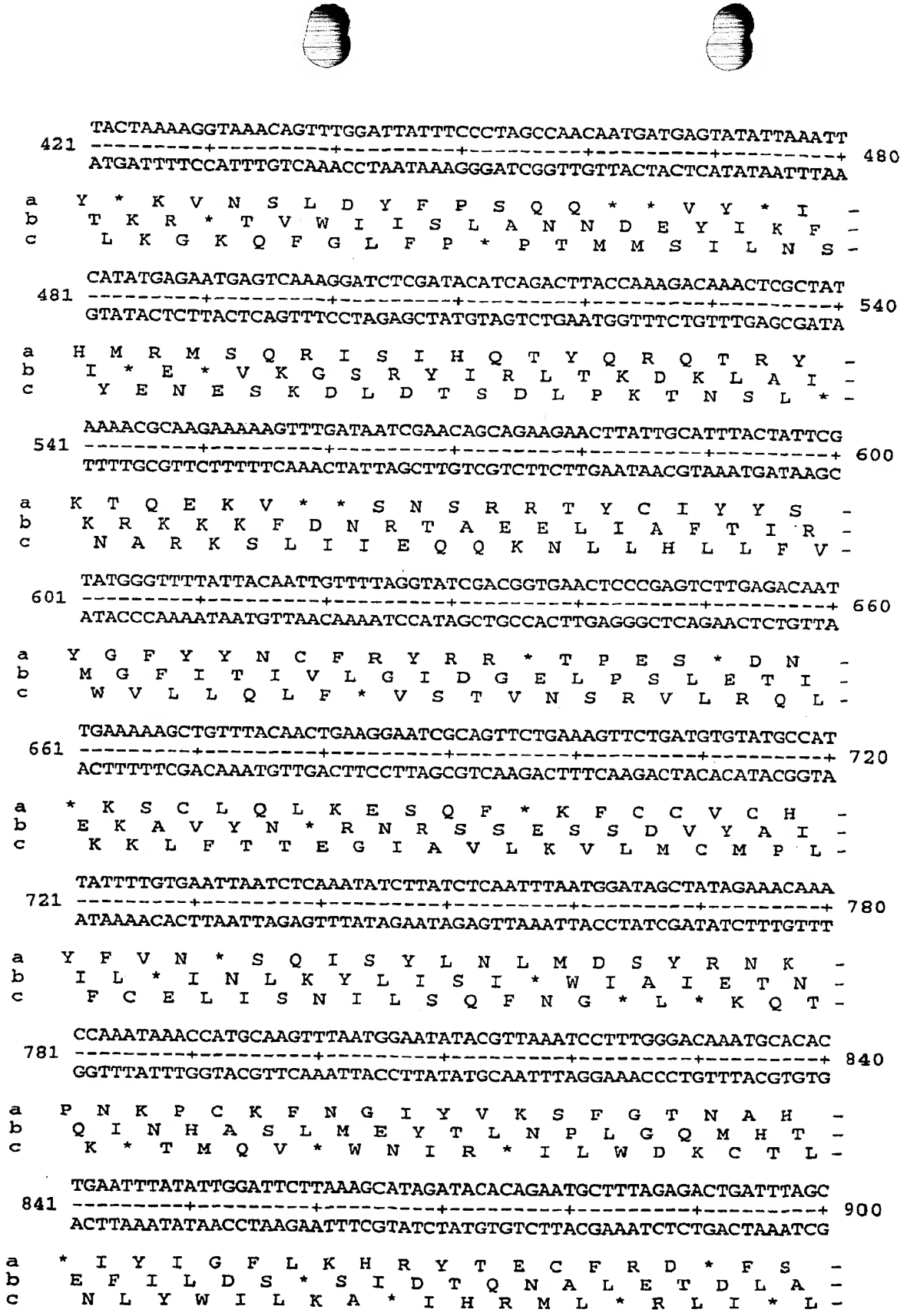


FIG. 35
(CONTINUED)

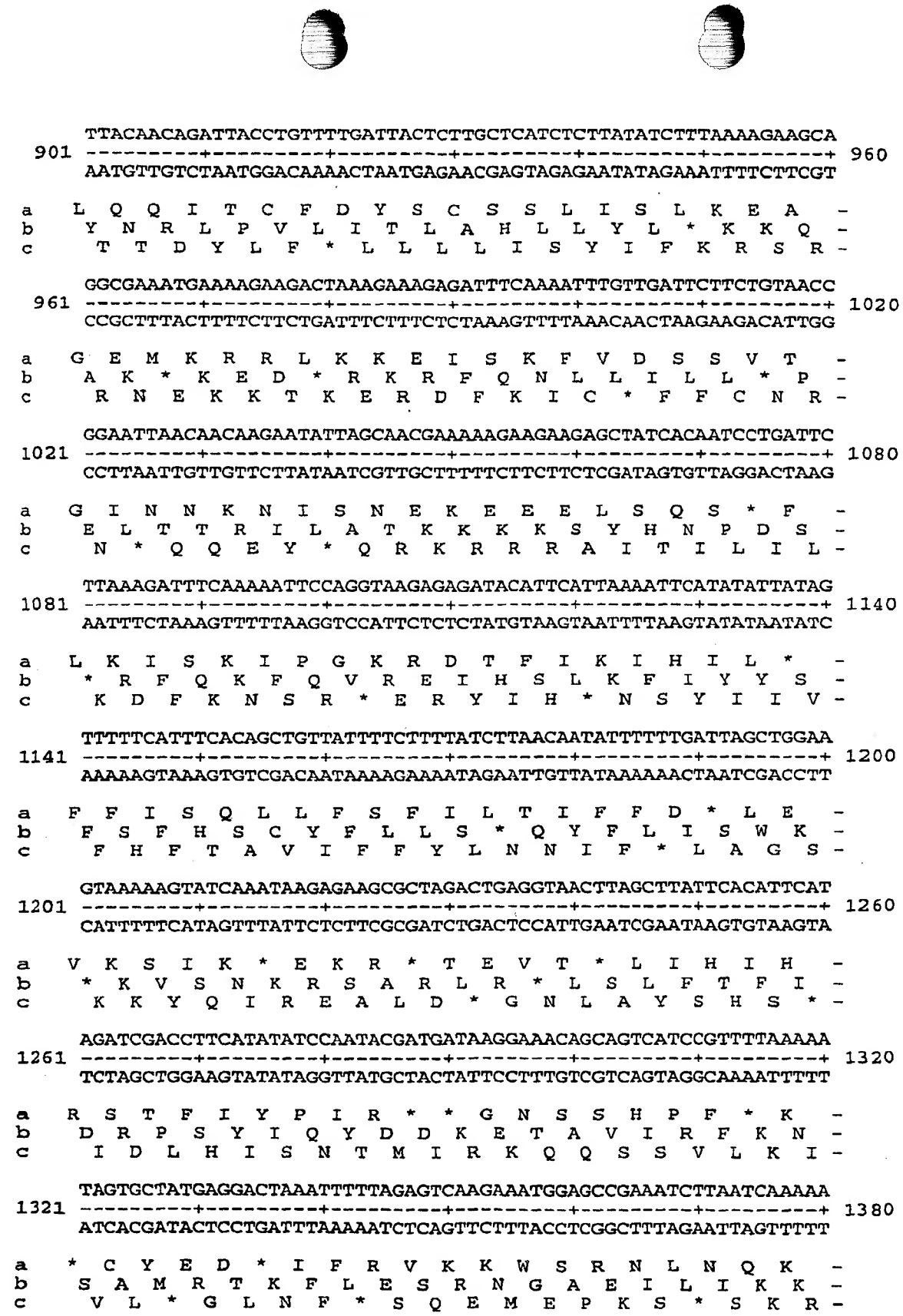


FIG. 35
(CONTINUED)

```
     GAATTGCGTCGATATTGCAAAAGAATCGAACTCTAAATCTTTCGTTAATAAGTATTACCA
1381 ---------+---------+---------+---------+---------+---------+ 1440
     CTTAACGCAGCTATAACGTTTTCTTAGCTTGAGATTTAGAAAGCAATTATTCATAATGGT

a    E  L  R  R  Y  C  K  R  I  E  L  *  I  F  R  *  *  V  L  P  -
b     N  C  V  D  I  A  K  E  S  N  S  K  S  F  V  N  K  Y  Y  Q  -
c      I  A  S  I  L  Q  K  N  R  T  L  N  L  S  L  I  S  I  T  N -

     ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAAGAAATAAA
1441 ---------+---------+---------+---------+---------+---------+ 1500
     TAGAACTAACTAACTTCTCTAACTGCTCCGTTGACGTGTCTTCTAGTAATTTCTTTATTT

a    I  L  I  D  C  R  D  C  R  G  N  C  T  E  D  H  *  R  N  K  -
b     S  *  L  I  E  E  I  D  E  A  T  A  Q  K  I  I  K  E  I  K  -
c      L  D  *  L  K  R  L  T  R  Q  L  H  R  R  S  L  K  K  *  S -

     GTAACTTTTATTAATTAGAGAATAAACTAAATTACTAATATAGAGATCAGCGATCTTCAA
1501 ---------+---------+---------+---------+---------+---------+ 1560
     CATTGAAAATAATTAATCTCTTATTTGATTTAATGATTATATCTCTAGTCGCTAGAAGTT

a    V  T  F  I  N  *  R  I  N  *  I  T  N  I  E  I  S  D  L  Q  -
b     *  L  L  L  I  R  E  *  T  K  L  L  I  *  R  S  A  I  F  N  -
c      N  F  Y  *  L  E  N  K  L  N  Y  *  Y  R  D  Q  R  S  S  I -

     TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAAAATACAAACCTTGGTCAAAAT
1561 ---------+---------+---------+---------+---------+---------+ 1620
     AACTGCTTTATTTTCGACTTGATTTCAATCTGTTATTTTTTATGTTTGGAACCAGTTTTA

a    L  T  K  *  K  L  N  *  S  *  T  I  K  N  T  N  L  G  Q  N  -
b     *  R  N  K  S  *  T  K  V  R  Q  *  K  I  Q  T  L  V  K  I  -
c      D  E  I  K  A  E  L  K  L  D  N  K  K  Y  K  P  W  S  K  Y -

     ATTGAGGAAGGAAAAGAAGACCAGTTAGCAAAAGAAAAAATAAGGCAATAAATAAAATGA
1621 ---------+---------+---------+---------+---------+---------+ 1680
     TAACTCCTTCCTTTTCTTCTGGTCAATCGTTTTCTTTTTTATTCCGTTATTTATTTTACT

a    I  E  E  G  K  E  D  Q  L  A  K  E  K  I  R  Q  *  I  K  *  -
b     L  R  K  E  K  K  T  S  *  Q  K  K  K  *  G  N  K  *  N  E  -
c      *  G  R  K  R  R  P  V  S  K  R  K  N  K  A  I  N  K  M  S -

     GTACAGAAGTGAAGAAATAAAAGATTTATTTTTTTCAATAATTTATTGAAAAGAGGGGTT
1681 ---------+---------+---------+---------+---------+---------+ 1740
     CATGTCTTCACTTCTTTATTTTCTAAATAAAAAAAGTTATTAAATAACTTTTCTCCCCAA

a    V  Q  K  *  R  N  K  R  F  I  F  F  N  N  L  L  K  R  G  V  -
b     Y  R  S  E  E  I  K  D  L  F  F  S  I  I  Y  *  K  E  G  F  -
c      T  E  V  K  K  *  K  I  Y  F  F  Q  *  F  I  E  K  R  G  F -

     TTGGGGTTTTGGGGTTTTGGGG
1741 ---------+---------+-- 1762
     AACCCCAAAACCCCAAAACCCC

a    L  G  F  W  G  F  G  -
b     W  G  F  G  V  L  G  -
c      G  V  L  G  F  W  -
```

FIG. 35
(CONTINUED)

Figure 36

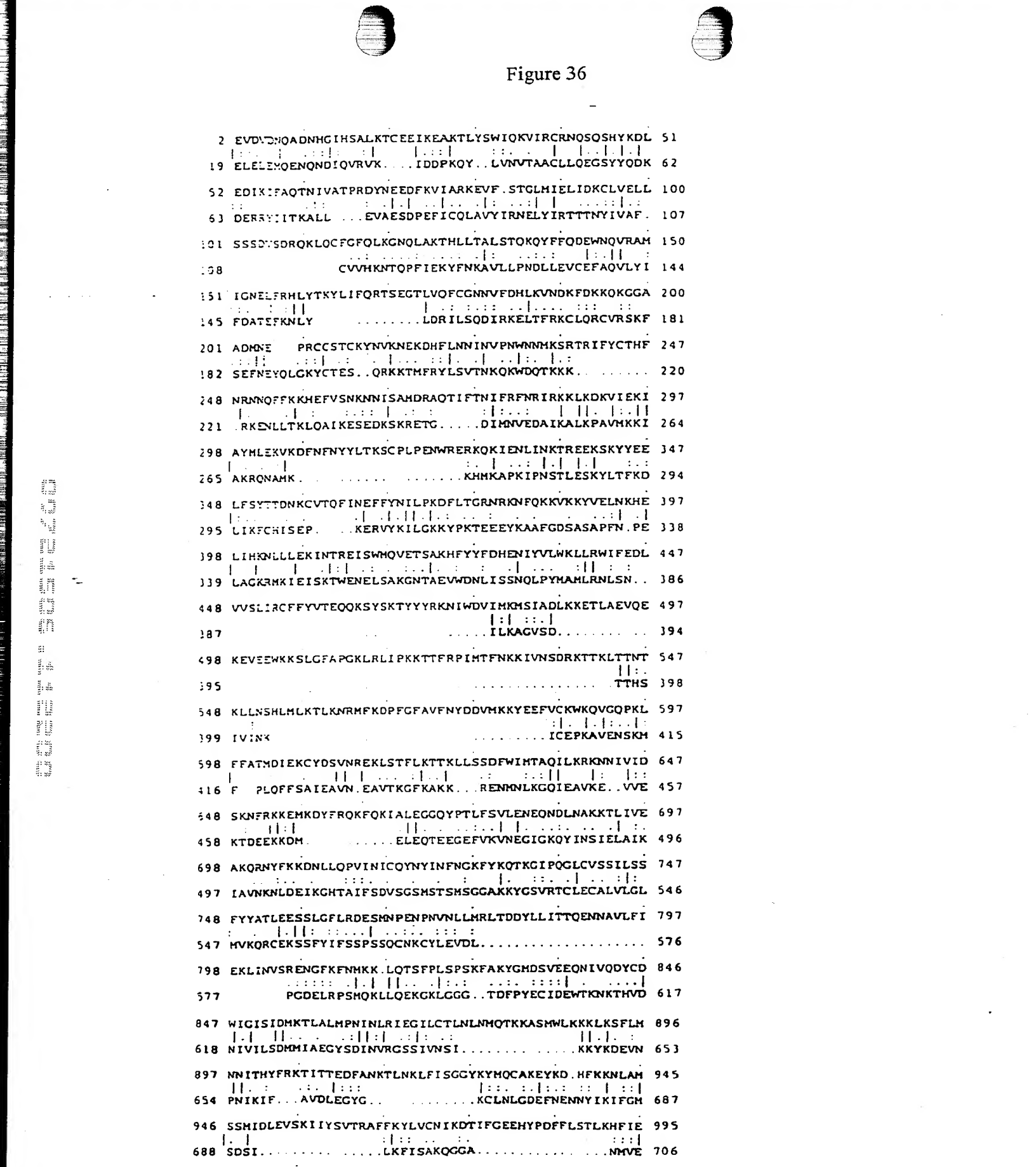


Figure 37

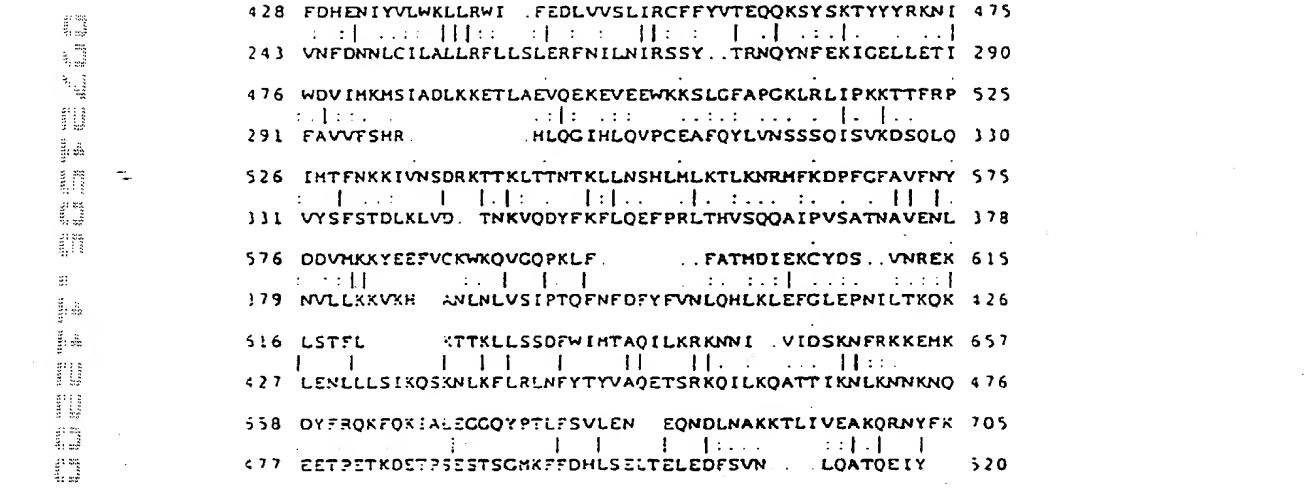


Figure 38

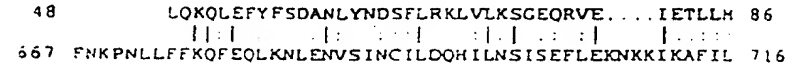

4 DIDLDDIENLLPNTFNKYSSSCSDKKGCKTLKSGSKSPSLTIPK 47
617 NVKSAKIESSSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPNNIQKNP 666
48 LQKQLEFYFSDANLYNDSFLRKLVLKSGEQRVE IETLLH 86
667 FNKPNLLFFKQFEQLKNLENVSINCILDQHILNSISEFLEKNKKIKAFIL 716

Figure 39

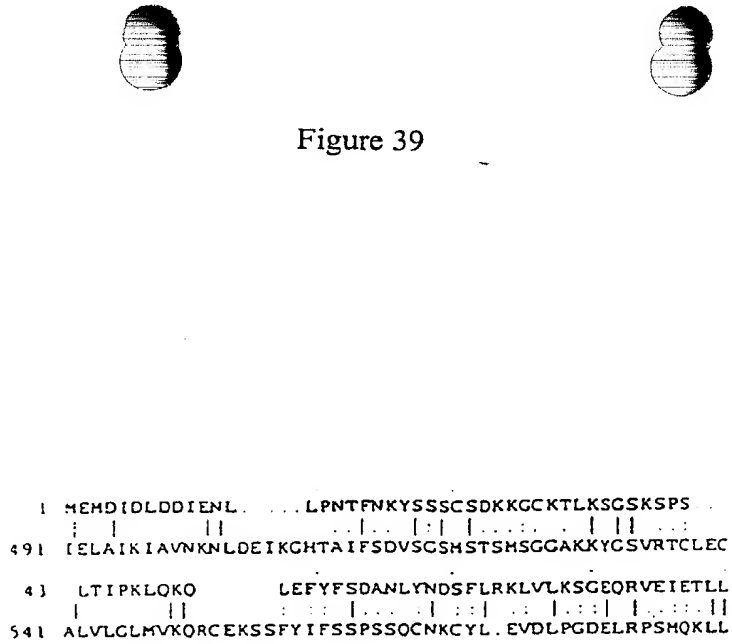

1 MEMDIDLDDIENL . . . LPNTFNKYSSSCSDKKGCKTLKSGSKSPS . 42
491 IELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGAKKYGSVRTCLEC 540
43 LTIPKLQKQ LEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLL 85
541 ALVLGLMVKQRCEKSSFYIFSSPSSQCNKCYL . EVDLPGDELRPSMQKLL 589

Figure 40

```
                               Motif A                                        Motif B
Consensus                 h--hDh---h--h                              h---+-QG---Sp
telomerase p123 GQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLL-100-KFYRQTKGIPQGLCVSSILSSFYYATLEESSLGPL
Dong (LINE)     KNRNLHCTYIDYKKAFDSIPHSWLIQVLEIYKIN- 28-RQIAIKKGIYQGDSLSPLWFCLALNPLSHQLHNDR
al S.c.(groupII)FGGSNWFREVDLKKCFDTISHDLIIKELKRYISD- 26-HVPVGPRVCVQGAPTSPALCNAVLLRLDRRLAGLA
HIV-RT          LKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIP-  7-GIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQN
L8543.12 YAM    VLPELYFMKFDVKSCYDSIPRMECMRILKDALKN- 68-KCYIREDGLFQGSSLSAPIVDLVYDDLLEFYSEFK

                  Motif C                                  Motif D          Motif E
Consensus           h--YhDDhhh                           Gh-h---K          h-hLGh-h
telomerase p123 -14-LMRLTDDYLLITTQENN-0-AVLFIEKLINVSRENGFKFNMKKLQT-23-QDYCDWIGISI
Dong (LINE)     -16-HLIYMDDIKLYAKNDKE-0-MKKLIDTTTIFSNDISMQFGLDKCKT-25-KCLYKYLGFQQ
al S.c.(groupII)-55-YVRYADDILIGVLGSKN-2-KIIKRDLNNFLNS.LGLTINEEKTLI- 4-ETPARFLGYNI
HIV-RT          - 4-IYQYMDDLYVGSHLEIG-1-HRTKIEELRQHLLRWGLTTPDKKHQK- 0-EPPFLWMGYEL
L8543.12 YAM    - 8-ILKLADDFLIISTDQQQ.... ...VINIKKLAMGGFQKYNAKANR-41-IRSKSSKGIFR
```

Figure 41

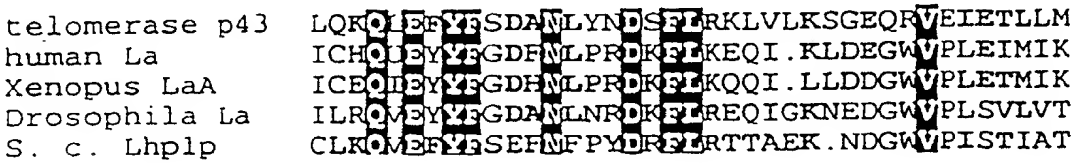

telomerase p43 LQKQLEFYESDANLYNDSFLRKLVLKSGEQRVEIETLLM
human La ICHQLEYYEGDFNLPRDKFLKEQI.KLDEGWVPLEIMIK
Xenopus LaA ICEQLEYYEGDHNLPRDKFLKQQI.LLDDGWVPLETMIK
Drosophila La ILRQMEYYEGDANLNRDKFLREQIGKNEDGWVPLSVLVT
S. c. Lhp1p CLKQMEFYESEFNFPYDRFLRTTAEK.NDGWVPISTIAT

Figure 42

```
   1 aactcattta attactaatt taatcaacaa gattgataaa aagcagtaaa taaaacccaa
  61 tagatttaat ttagaaagta tcaattgaaa aatggaaatt gaaaacaact aagcacaata
 121 gccaaaagcc gaaaaattgt ggtgggaact tgaattagag atgcaagaaa accaaaatga
 181 tatataagtt agggttaaga ttgacgatcc taagcaatat ctcgtgaacg tcactgcagc
 241 atgtttgttg taggaaggta gttactacta agataaagat gaaagaagat atatcatcac
 301 taaagcactt cttgaggtgg ctgagtctga tcctgagttc atctgctagt tggcagtcta
 361 catccgtaat gaactttaca tcagaactac cactaactac attgtagcat tttgtgttgt
 421 ccacaagaat actcaaccat tcatcgaaaa gtacttcaac aaagcagtac ttttgcctaa
 481 tgacttactg gaagtctgtg aatttgcata ggttctctat atttttgatg caactgaatt
 541 caaaaatttg tatcttgata ggatactttc ataagatatt cgtaaggaac tcactttccg
 601 taagtgttta caaagatgcg tcagaagcaa gttttctgaa ttcaacgaat actaacttgg
 661 taagtattgc actgaatcct aacgtaagaa aacaatgttc cgttacctct cagttaccaa
 721 caagtaaaag tgggattaaa ctaagaagaa gagaaaagag aatctcttaa ccaaacttta
 781 ggcaataaag gaatctgaag ataagtccaa gagagaaact ggagacataa tgaacgttga
 841 agatgcaatc aaggctttaa aaccagcagt tatgaagaaa atagccaaga gatagaatgc
 901 catgaagaaa cacatgaagg cacctaaaat tcctaactct accttggaat caaagtactt
 961 gaccttcaag gatctcatta agttctgcca tatttctgag cctaaagaaa gagtctataa
1021 gatccttggt aaaaaatacc ctaagaccga agaggaatac aaagcagcct ttggtgattc
1081 tgcatctgca cccttcaatc ctgaattggc tggaaagcgt atgaagattg aaatctctaa
1141 aacatgggaa aatgaactca gtgcaaaagg caacactgct gaggtttggg ataatttaat
1201 ttcaagcaat taactcccat atatggccat gttacgtaac ttgtctaaca tcttaaaagc
1261 cggtgtttca gatactacac actctattgt gatcaacaag atttgtgagc ccaaggccgt
1321 tgagaactcc aagatgttcc ctcttcaatt ctttagtgcc attgaagctg ttaatgaagc
1381 agttactaag ggattcaagg ccaagaagag agaaaatatg aatcttaaag gtcaaatcga
1441 agcagtaaag gaagttgttg aaaaaaccga tgaagagaag aaagatatgg agttggagta
1501 aaccgaagaa ggagaatttg ttaaagtcaa cgaaggaatt ggcaagcaat acattaactc
1561 cattgaactt gcaatcaaga tagcagttaa caagaattta gatgaaatca aaggacacac
1621 tgcaatcttc tctgatgttt ctggttctat gagtacctca atgtcaggtg gagccaagaa
1681 gtatggttcc gttcgtactt gtctcgagtg tgcattagtc cttggtttga tggtaaaata
1741 acgttgtgaa aagtcctcat tctacatctt cagttcacct agttctcaat gcaataagtg
1801 ttacttagaa gttgatctcc ctggagacga actccgtcct tctatgtaaa aacttttgca
1861 agagaaagga aaacttggtg gtggtactga tttccctat gagtgcattg atgaatggac
1921 aaagaataaa actcacgtag acaatatcgt tattttgtct gatatgatga ttgcagaagg
1981 atattcagat atcaatgtta gaggcagttc cattgttaac agcatcaaaa agtacaagga
2041 tgaagtaaat cctaacatta aaatctttgc agttgactta gaaggttacg gaaagtgcct
2101 taatctaggt gatgagttca atgaaaacaa ctacatcaag atattcggta tgagcgattc
2161 aatcttaaag ttcatttcag ccaagcaagg aggagcaaat atggtcgaag ttatcaaaaa
2221 ctttgccctt caaaaaatag gacaaaagtg agtttcttga gattcttcta taacaaaaat
2281 ctcaccccac tttttgttt tattgcatag ccattatgaa atttaaatta ttatctattt
2341 atttaagtta cttacatagt ttatgtatcg cagtctatta gcctattcaa atgattctgc
2401 aaagaacaaa aaagattaaa a
```

Figure 43

MEIENNQAQQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL

VNVTAACLLQEGSYYQDKDERRYIITKALLEVAESDPEFICQLAVYIRNELYIRTTTN

YIVAFCVVHKNTQPFIEKYFNKAVLLPNDLLEVCEFAQVLYIFDATEFKNLYLDRILS

QDIRKELTFRKCLQRCVRSKFSEFNEYQLGKYCTESQRKKTMFRYLSVTNKQKWDQTK

KKRKENLLTKLQAIKESEDKSKRETGDIMNVEDAIKALKPAVMKKIAKRQNAMKKHMK

APKIPNSTLESKYLTFKDLIKFCHISEPKERVYKILGKKYPKTEEEYKAAFGDSASAP

FNPELAGKRMKIEISKTWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNILKAGV

SDTTHSIVINKICEPKAVENSKMFPLQFFSAIEAVNEAVTKGFKAKKRENMNLKGQIE

AVKEVVEKTDEEKKDMELEQTEEGEFVKVNEGIGKQYINSIELAIKIAVNKNLDEIKG

HTAIFSDVSGSMSTSMSGGAKKYGSVRTCLECALVLGLMVKQRCEKSSFYIFSSPSSQ

CNKCYLEVDLPGDELRPSMQKLLQEKGKLGGGTDFPYECIDEWTKNKTHVDNIVILSD

MMIAEGYSDINVRGSSIVNSIKKYKDEVNPNIKIFAVDLEGYGKCLNLGDEFNENNYI

KIFGMSDSILKFISAKQGGANMVEVIKNFALQKIGQK

Figure 44

1 tcaatactat taattaataa ataaaaaaaa gcaaactaca aagaaaatgt caaggcgtaa
61 ctaaaaaaag ccataggctc ctataggcaa tgaaacaaat cttgattttg tattacaaaa
121 tctagaagtt tacaaaagcc agattgagca ttataagacc tagtagtaat agatcaaaga
181 ggaggatctc aagcttttaa agttcaaaaa ttaagattag gatggaaact ctggcaacga
241 tgatgatgat gaagaaaaca actcaaataa ataataagaa ttattaagga gagtcaatta
301 gattaagtag caagtttaat tgataaaaaa agttggttct aaggtagaga aagatttgaa
361 tttgaacgaa gatgaaaaca aaaagaatgg actttctgaa tagcaagtga aagaagagta
421 attaagaacg attactgaag aataggttaa gtattaaaat ttagtattta acatggacta
481 ccagttagat ttaaatgaga gtggtggcca tagaagacac agaagagaaa cagattatga
541 tactgaaaaa tggtttgaaa tatctcatga ccaaaaaaat tatgtatcaa tttacgccaa
601 ctaaaagaca tcatattgtt ggtggcttaa agattatttt aataaaaaca attatgatca
661 tcttaatgta agcattaaca gactagaaac tgaagccgaa ttctatgcct ttgatgattt
721 ttcacaaaca atcaaactta ctaataattc ttactagact gttaacatag acgttaattt
781 tgataataat ctctgtatac tcgcattgct tagatttttа ttatcactag aaagattcaa
841 tattttgaat ataagatctt cttatacaag aaattaatat aattttgaga aaattggtga
901 gctacttgaa actatcttcg cagttgtctt ttctcatcgc cacttacaag gcattcattt
961 acaagttcct tgcgaagcgt tctaatattt agttaactcc tcatcataaa ttagcgttaa
1021 agatagctaa ttataggtat actctttctc tacagactta aaattagttg acactaacaa
1081 agtccaagat tattttaagt tcttataaga attcccctcgt ttgactcatg taagctagta
1141 ggctatccca gttagtgcta ctaacgctgt agagaacctc aatgttttac ttaaaaaggt
1201 caagcatgct aatcttaatt tagtttctat ccctacctaa ttcaattttg atttctactt
1261 tgttaattta taacatttga aattagagtt tggattagaa ccaaatattt tgacaaaaca
1321 aaagcttgaa aatctacttt tgagtataaa ataatcaaaa aatcttaaat ttttaagatt
1381 aaacttttac acctacgttg cttaagaaac ctccagaaaa cagatattaa aacaagctac
1441 aacaatcaaa aatctcaaaa acaataaaaa tcaagaagaa actcctgaaa ctaaagatga
1501 aactccaagc gaaagcacaa gtggtatgaa attttttgat catctttctg aattaaccga
1561 gcttgaagat ttcagcgtta acttgtaagc tacccaagaa atttatgata gcttgcacaa
1621 acttttgatt agatcaacaa atttaaagaa gttcaaatta agttacaaat atgaaatgga
1681 aaagagtaaa atggatacat tcatagatct taagaatatt tatgaaacct taaacaatct
1741 taaaagatgc tctgttaata tatcaaatcc tcatggaaac atttcttatg aactgacaaa
1801 taaagattct acttttttata aatttaagct gaccttaaac taagaattat aacacgctaa
1861 gtatacttt aagtagaacg aatttaatt taataacgtt aaaagtgcaa aaattgaatc
1921 ttccttcatta gaaagcttag aagatattga tagtctttgc aaatctattg cttcttgtaa
1981 aaatttacaa aatgttaata ttatcgccag tttgctctat cccaacaata tttagaaaaa
2041 tccttttcaat aagcccaatc ttctatttttt caagcaattt gaataattga aaaatttgga
2101 aaatgtatct atcaactgta ttcttgatca gcatatactt aattctattt cagaattctt
2161 agaaaagaat aaaaaaataa aagcattcat tttgaaaaga tattatttat tacaatatta
2221 tcttgattat actaaattat ttaaaacact tcaatagtta cctgaattaa attaagttta
2281 cattaattag caattagaag aattgactgt gagtgaagta cataagtaag tatgggaaaa
2341 ccacaagcaa aaagctttct atgaaccatt atgtgagttt atcaaagaat catcctaaac
2401 cctttagcta atagattttg accaaaacac tgtaagtgat gactctatta aaaagatttt
2461 agaatctata tctgagtcta agtatcatca ttatttgaga ttgaacccta gttaatctag
2521 cagtttaatt aaatctgaaa acgaagaaat ttaagaactt ctcaaagctt gcgacgaaaa
2581 aggtgttta gtaaaagcat actataaatt ccctctatgt ttaccaactg gtacttatta
2641 cgattacaat tcagatagat ggtgattaat taaatattag tttaaataaa tattaaatat
2701 tgaatatttc tttgcttatt atttgaataa tacatacaat agtcatttтt agtgttttga
2761 atatatttta gttatttaat tcattatttt aagtaaataa ttatttttca atcatttttt
2821 aaaaaatcg

Figure 45

MSRRNQKKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQQIK

EEDLKLLKFKNQDQDGNSGNDDDDEENNSNKQQELLRRVNQIKQQVQLIKKVGSKVEK

DLNLNEDENKKNGLSEQQVKEEQLRTITEEQVKYQNLVFNMDYQLDLNESGGHRRHRR

ETDYDTEKWFEISHDQKNYVSIYANQKTSYCWWLKDYFNKNNYDHLNVSINRLETEAE

FYAFDDFSQTIKLTNNSYQTVNIDVNFDNNLCILALLRFLLSLERFNILNIRSSYTRN

QYNFEKIGELLETIFAVVFSHRHLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQVYSF

STDLKLVDTNKVQDYFKFLQEFPRLTHVSQQAIPVSATNAVENLNVLLKKVKHANLNL

VSIPTQFNFDFYFVNLQHLKLEFGLEPNILTKQKLENLLLSIKQSKNLKFLRLNFYTY

VAQETSRKQILKQATTIKNLKNNKNQEETPETKDETPSESTSGMKFFDHLSELTELED

FSVNLQATQEIYDSLHKLLIRSTNLKKFKLSYKYEMEKSKMDTFIDLKNIYETLNNLK

RCSVNISNPHGNISYELTNKDSTFYKFKLTLNQELQHAKYTFKQNEFQFNNVKSAKIE

SSSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPNNIQKNPFNKPNLLFFKQFEQLK

NLENVSINCILDQHILNSISEFLEKNKKIKAFILKRYYLLQYYLDYTKLFKTLQQLPE

LNQVYINQQLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESSQTLQLIDFDQNTVSD

DSIKKILESISESKYHHYLRLNPSQSSSLIKSENEEIQELLKACDEKGVLVKAYYKFP

LCLPTGTYYDYNSDRW

Figure 46

MKILFEFIQDKLDIDLQTNSTYKENLKCGHFNGLDEILTTCFAL

PNSRKIALPCLPGDLSHKAVIDHCIIYLLTGELYNNVLTFGYKIARNEDVNNSLFCHS

ANVNVTLLKGAAWKMFHSLVGTYAFVDLLINYTVIQFNGQFFTQIVGNRCNEPHLPPK

WVQRSSSSSATAAQIKQLTEPVTNKQFLHKLNINSSSFFPYSKILPSSSSIKKLTDLR

EAIFPTNLVKIPQRLKVRINLTLQKLLKRHKRLNYVSILNSICPPLEGTVLDLSHLSR

QSPKERVLKFIIVILQKLLPQEMFGSKKNKGKIIKNLNLLLSLPLNGYLPFDSLLKKL

RLKDFRWLFISDIWFTKHNFENLNQLAICFISWLFRQLIPKIIQTFFYCTEISSTVTI

VYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLSNFNHSKMRIIPKKSNNEFR

IIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKE

FKQRLLKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFN

TNTGVLKLFNVVVNASRVPKPYELYIDNVRTVHLSNQDVINVVEMEIFKTALWVEDKCY

IREDGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPSQDTLILKLADDFLIISTDQQQV

INIKKLAMGGFQKYNAKANRDKILAVSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMN

NFHIRSKSSKGIFRSLIALFNTRISYKTIDTNLNSTNTVLMQIDHVVKNISECYKSAF

KDLSINVTQNMQFHSFLQRIIEMTVSGCPITKCDPLIEYEVRFTILNGFLESLSSNTS

KFKDNIILLRKEIQHLQAYIYIYIHIVN

Figure 47

Oxytricha LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
Euplotes LCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLT

Figure 48

ATTTATACTCATGAAAATCTTATTCGAGTTCATTCAAGACAAGCTTGACATTGATCTACA
GACCAACAGTACTTACAAAGAAAATTTAAAATGTGGTCACTTCAATGGCCTCGATGAAAT
TCTAACTACGTGTTTCGCACTACCAAATTCAAGAAAAATAGCATTACCATGCCTTCCTGG
TGACTTAAGCCACAAAGCAGTCATTGATCACTGCATCATTTACCTGTTGACGGGCGAATT
ATACAACAACGTACTAACATTTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAATAG
TCTTTTTTGCCATTCTGCAAATGTTAACGTTACGTTACTGAAAGGCGCTGCTTGGAAAAT
GTTCCACAGTTTGGTCGGTACATACGCATTCGTTGATTTATTGATCAATTATACAGTAAT
TCAATTTAATGGGCAGTTTTTCACTCAAATCGTGGGTAACAGATGTAACGAACCTCATCT
GCCGCCCAAATGGGTCCAACGATCATCCTCATCATCCGCAACTGCTGCGCAAATCAAACA
ACTTACAGAACCAGTGACAAATAAACAATTCTTACACAAGCTCAATATAAATTCCTCTTC
TTTTTTTCCTTATAGCAAGATCCTTCCTTCATCATCATCTATCAAAAAGCTAACTGACTT
GAGAGAAGCTATTTTTCCCACAAATTTGGTTAAAATTCCTCAGAGACTAAAGGTACGAAT
TAATTTGACGCTGCAAAAGCTATTAAAGAGACATAAGCGTTTGAATTACGTTTCTATTTT
GAATAGTATTTGCCCACCATTGGAAGGGACCGTATTGGACTTGTCGCATTTGAGTAGGCA
ATCACCAAAGGAACGAGTCTTGAAATTTATCATTGTTATTTTACAGAAGTTATTACCCCA
AGAAATGTTTGGCTCAAAGAAAAATAAAGGAAAAATTATCAAGAATCTAAATCTTTTATT
AAGTTTACCCTTAAATGGCTATTTACCATTTGATAGTTTGTTGAAAAAGTTAAGATTAAA
GGATTTTCGGTGGTTGTTCATTTCTGATATTTGGTTCACCAAGCACAATTTTGAAAACTT
GAATCAATTGGCGATTTGTTTCATTTCCTGGCTATTTAGACAACTAATTCCCAAAATTAT
ACAGACTTTTTTTTACTGCACCGAAATATCTTCTACAGTGACAATTGTTTACTTTAGACA
TGATACTTGGAATAAACTTATCACCCCTTTTATCGTAGAATATTTTAAGACGTACTTAGT
CGAAAACAACGTATGTAGAAACCATAATAGTTACACGTTGTCCAATTTCAATCATAGCAA
AATGAGGATTATACCAAAAAAAAGTAATAATGAGTTCAGGATTATTGCCATCCCATGCAG
AGGGGCAGACGAAGAAGAATTCACAATTTATAAGGAGAATCACAAAAATGCTATCCAGCC
CACTCAAAAAATTTTAGAATACCTAAGAAACAAAAGGCCGACTAGTTTTACTAAAATATA
TTCTCCAACGCAAATAGCTGACCGTATCAAAGAATTTAAGCAGAGACTTTTAAAGAAATT
TAATAATGTCTTACCAGAGCTTTATTTCATGAAATTTGATGTCAAATCTTGCTATGATTC
CATACCAAGGATGGAATGTATGAGGATACTCAAGGATGCGCTAAAAAATGAAAATGGGTT
TTTCGTTAGATCTCAATATTTCTTCAATACCAATACAGGTGTATTGAAGTTATTTAATGT
TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTATACATAGATAATGTGAGGACGGT
TCATTTATCAAATCAGGATGTTATAAACGTTGTAGAGATGGAAATATTTAAAACAGCTTT
GTGGGTTGAAGATAAGTGCTACATTAGAGAAGATGGTCTTTTTCAGGGCTCTAGTTTATC
TGCTCCGATCGTTGATTTGGTGTATGACGATCTTCTGGAGTTTTATAGCGAGTTTAAAGC
CAGTCCTAGCCAGGACACATTAATTTTAAAACTGGCTGACGATTTCCTTATAATATCAAC
AGACCAACAGCAAGTGATCAATATCAAAAAGCTTGCCATGGGCGGATTTCAAAAATATAA
TGCGAAAGCCAATAGAGACAAAATTTTAGCCGTAAGCTCCCAATCAGATGATGATACGGT
TATTCAATTTTGTGCAATGCACATATTTGTTAAAGAATTGGAAGTTTGGAAACATTCAAG
CACAATGAATAATTTCCATATCCGTTCGAAATCTAGTAAAGGGATATTTCGAAGTTTAAT
AGCGCTGTTTAACACTAGAATCTCTTATAAAACAATTGACACAAATTTAAATTCAACAAA
CACCGTTCTCATGCAAATTGATCATGTTGTAAAGAACATTTCGGAATGTTATAAATCTGC
TTTTAAGGATCTATCAATTAATGTTACGCAAAATATGCAATTTCATTCGTTCTTACAACG
CATCATTGAAATGACAGTCAGCGGTTGTCCAATTACGAAATGTGATCCTTTAATCGAGTA
TGAGGTACGATTCACCATATTGAATGGATTTTTGGAAAGCCTATCTTCAAACACATCAAA
ATTTAAAGATAATATCATTCTTTTGAGAAAGGAAATTCAACACTTGCAAGC

Figure 49

AKFLHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKR
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKR
AERLTSRVKALFSVLNYERA

Figure 50

GCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTC
TTTCTTTTATGTCACGGAGACCACGTTTCAAAAGAACAGGCTCTTTTTCTACC
GGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAA
GAGGGTGCAGCTGCGGGACGTGTCGGAAGCAGAGGTCAGGCAGCATCGGGA
AGCCAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACG
GGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTTCCG
CAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTCAGC
GTGCTCAACTACGAGCGGGCGCG

Figure 51

MTEHHTPKSRILRFLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLRSDVQTSFSIFLHSTVVGF
DSKPDEGVQFSSPKCSQSELIANVVKQMFDESFERRRNLLMKGFSMNHEDFRAMHVNGVQNDLV
STFPNYLISILESKNWQLLLEIIGSDAMHYLLSKGSIFEALPNDNYLQISGIPLFKNNVFEETVSKKRK
RTIETSITQNKSARKEVSWNSISISRFSIFYRSSYKKFKQDLYFNLHSICDRNTVHMWLQWIFPRQFG
LINAFQVKQLHKVIPLVSQSTVVPKRLLKVYPLIEQTAKRLHRISLSKVYNHYCPYIDTHDDEKILS
YSLKPNQVFAFLRSILVRVFPKLIWGNQRIFEIILKDLETFLKLSRYESFSLHYLMSNIKISEIEWLVL
GKRSNAKMCLSDFEKRKQIFAEFIYWLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFIT
SMKMEAFEKINENNVRMDTQKTTLPPAVIRLLPKKNTFRLITNLRKRFLIKMGSNKKMLVSTNQT
LRPVASILKHLINEESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDIKSCYDRIKQDLMFR
IVKKKLKDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPFEKVVQLLSMKTSDTLFVDFVDY
WTKSSSEIFKMLKEHLSGHIVKIGNSQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL
LRVVDDFLFITVNKKDAKKFLNLSLRGFEKHNFSTSLEKTVINFENSNGIINNTFFNESKKRMPFFG
FSVNMRSLDTLLACPKIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCNI
YRLGYSMCMRAQAYLKRMKDIFIPQRMFITDLLNVIGRKIWKKLAEILGYTSRRFLSSAEVKWLFC
LGMRDGLKPSFKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVLFLHRRIAD

Figure 52

ggtaccgatttacttccttcttcataagctaattgcttcctcgaacgctcctaaatctctggaaatatttttacaagaactcaataacaataccaagtcaaattccaatatgaagg
tgttattagtgatcgataatattctatttttatcggtcgttaccaagtataaggacaaaaagaacaacttccttccccctaaagacttttacttattaatttacttttcaaatatatttcg
ggttcgcttacttttaatcgtggtactgttttagctgctacttctagccaaccgcgtgtttctaccccgtcattggatatagctcttggagtagctcacagaaatccttacaaatctt
ctgatgagactatattagattcattacagtccgtgcatattcttaacatggagccttacactttagatgagtcacgtcgcatgatggagtatttggtatcatccaacgtttgccttg
aaaaggttgataattattgcaaaatcatgtccttagtggtggtaatccgcgaaagtttttgatgcttgcacacgtctagcatgattgagatattcaaaaatttctatccactacaa
ctccttttaacgcggttttatttttctatttctattctcatgttgttccaaatatgtatcatctcgtattaggcttttttccgttttactccttggaatcgtaccttttcactattcccccctaatg
aataatctaaattagttcgcttataattgatagtagtagaaagattggtgattctactcgtgtaatgttattagtttaaagatacttgcaaaacatttattagctatcattatataaaa
aaaatcctataattataaatattaatcaatatttgcggtcactatttatttaaaacgttatgatcagtaggacacttgcatatatatagttatgcttaatggttacttgtaacttgcAT
GACCGAACACCATACCCCCAAAAGCAGGATTCTTCGCTTTCTAGAGAATCAATATGTATACCTATGTA
CCTTAAATGATTATGTACAACTTGTTTTGAGAGGGTCGCCGGCAAGCTCGTATAGCAATATATGCGAA
CGCTTGAGAAGCGATGTACAAACGTCCTTTTCTATTTTTCTTCATTCGACTGTAGTCGGCTTCGACAGT
AAGCCAGATGAAGGTGTTCAATTTTCTTCTCCAAAATGCTCACAGTCAGAGgtatatatattttgttttgattttttctattcg
ggatagctaatatatgggcagCTAATAGCGAATGTTGTAAAACAGATGTTCGATGAAAGTTTTGAGCGTCGAAGGA
ATCTACTGATGAAAGGGTTTTCCATGgtaaggtattctaattgtgaaatatttacctgcaattactgtttcaaagagattgtatttaaccgataaagAA
TCATGAAGATTTTCGAGCCATGCATGTAAACGGAGTACAAAATGATCTCGTTTCTACTTTTCCTAATTA
CCTTATATCTATACTTGAGTCAAAAAATTGGCAACTTTTGTTAGAAATgtaaatacctggttaagatgttgcgcacttgaaca
agactgacaagtatagTATCGGCAGTGATGCCATGCATTACTTATTATCCAAAGGAAGTATTTTTGAGGCTCTTC
CAAATGACAATTACCTTCAGATTTCTGGCATACCACTTTTTAAAAATAATGTGTTTGAGGAAACTGTGT
CAAAAAAAAGAAAGCGAACCATTGAAACATCCATTACTCAAAATAAAAGCGCCCGCAAAGAAGTTTC
CTGGAATAGCATTTCAATTAGTAGGTTTAGCATTTTTTACAGGTCATCCTATAAGAAGTTTAAGCAAGgt
aactaatactgttatccttcataactaattttagATCTATATTTTAACTTACACTCTATTTGTGATCGGAACACAGTACACATG
TGGCTTCAATGGATTTTTCCAAGGCAATTTGGACTTATAAACGCATTTCAAGTGAAGCAATTGCACAA
AGTGATTCCACTGGTATCACAGAGTACAGTTGTGCCCAAACGTCTCCTAAAGGTATACCCTTTAATTGA
ACAAACAGCAAAGCGACTCCATCGTATTTCTCTATCAAAAGTTTACAACCATTATTGCCCATATATTGA
CACCCACGATGATGAAAAAATCCTTAGTTATTCCTTAAAGCCGAACCAGGTGTTTGCGTTTCTTCGATC
CATTCTTGTTCGAGTGTTTCCTAAATTAATCTGGGGTAACCAAAGGATATTTGAGATAATATTAAAAGg
tattgtataaaatttattaccactaacgattttaccagACCTCGAAACTTTCTTGAAATTATCGAGATACGAGTCTTTTAGTTTAC
ATTATTTAATGAGTAACATAAAGgtaatatgccaaatttttttaccattaattaacaatcagATTTCAGAAATTGAATGGCTAGT
CCTTGGAAAAAGGTCAAATGCGAAAATGTGCTTAAGTGATTTTGAGAAACGCAAGCAAATATTTGCGG
AATTCATCTACTGGCTATACAATTCGTTTATAATACCTATTTTACAATCTTTTTTTTATATCACTGAATC
AAGTGATTTACGAAATCGAACTGTTTATTTTAGAAAAGATATTTGGAAACTCTTGTGCCGACCCTTTAT
TACATCAATGAAAATGGAAGCGTTTGAAAAAATAAACGAGgtatttaaagtatttttgcaaaaagctaatatttcagAACAA
TGTTAGGATGGATACTCAGAAAACTACTTTGCCTCCAGCAGTTATTCGTCTATTACCTAAGAAGAATAC
CTTTCGTCTCATTACGAATTTAAGAAAAAGATTCTTAATAAAGgtattaatttttggtcatcaatgtacttacttctaatctattattag
cagATGGGTTCAAACAAAAAAATGTTAGTCAGTACGAACCAAACTTTACGACCTGTGGCATCGATACTG
AAACATTTAATCAATGAAGAAAGTAGTGGTATTCCATTTAACTTGGAGGTTTACATGAAGCTTCTTACT
TTTAAGAAGGATCTTCTTAAGCACCGAATGTTTGGgtaattatataatgcgcgattcctcattattaatttgcagGCGTAAGAAG
TATTTTGTACGGATAGATATAAAATCCTGTTATGATCGAATAAAGCAAGATTTGATGTTTCGGATTGTT
AAAAAGAAACTCAAGGATCCCGAATTTGTAATTCGAAAGTATGCAACCATACATGCAACAAGTGACCG
AGCTACAAAAAACTTTGTTAGTGAGGCGTTTTCCTATTgtaagttattttttcattggaatttttaacaaattcttttttagTTGATAT
GGTGCCTTTTGAAAAAGTCGTGCAGTTACTTTCTATGAAAACATCAGATACTTTGTTTGTTGATTTTGT
GGATTATTGGACCAAAAGTTCTTCTGAAATTTTTAAAATGCTCAAGGAACATCTCTCTGGACACATTGT
TAAGgtataccaattgttgaattgtaataacactaatgaaactagATAGGAAATTCTCAATACCTTCAAAAAGTTGGTATCCCTC
AGGGCTCAATTCTGTCATCTTTTTTGTGTCATTTCTATATGGAAGATTTGATTGATGAATACCTATCGTT
TACGAAAAAGAAAGGATCAGTGTTGTTACGAGTAGTCGACGATTTCCTCTTTATAACAGTTAATAAAA
AGGATGCAAAAAAATTTTTGAATTTATCTTTAAGAGgtgagttgctgtcattcctaagttctaaccgttgaagGATTTGAGAA
ACACAATTTTTCTACGAGCCTGGAGAAAACAGTAATAAACTTTGAAAATAGTAATGGGATAATAAACA
ATACTTTTTTTTAATGAAAGCAAGAAAAGAATGCCATTCTTCGGTTTCTCTGTGAACATGAGGTCTCTTG
ATACATTGTTAGCATGTCCTAAAATTGATGAAGCCTTATTTAACTCTACATCTGTAGAGCTGACGAAAC
ATATGGGGAAATCTTTTTTTTTACAAAATTCTAAGgtatactgtgtaactgaataatagctgacaaataatcagATCGAGCCTTGC
ATCCTTTGCACAAGTATTTATTGACATTACCCACAATTCAAAATTCAATTCTTGCTGCAATATATATAG
GCTAGGATACTCTATGTGTATGAGAGCACAAGCATACTTAAAAAGGATGAAGGATATATTTATTCCCC
AAAGAATGTTCATAACGGgtgagtacttatttaactagaaaagtcattaattaaccttagATCTTTTGAATGTTATTGGAAGAAAA
ATTTGGAAAAAGTTGGCCGAAATATTAGGATATACGAGTAGGCGTTTCTTGTCCTCTGCAGAAGTCAA

ATGgtacgtgtcggtctcgagacttcagcaatattgacacatcagGCTTTTTTGTCTTGGAATGAGAGATGGTTTGAAACCCTCTT
TCAAATATCATCCATGCTTCGAACAGCTAATATACCAATTTCAGTCATTGACTGATCTTATCAAGCCGC
TAAGACCAGTTTTGCGACAGGTGTTATTTTTACATAGAAGAATAGCTGATTAAtgtcatttcaattattatatacatcctt
tattactggtgtcttaaacaatattattactaagtatagctgacccccaaagcaagcatactataggatttctagtaaagtaaaattaatctcgttattagttttgattgacttgtcttt
atccttatactttttaagaaagattgacagtggttgctgactactgcccacatgcccattaaacgggagtggttaaacattaaaagtaatacatgaggctaatctcctttcatttag
aataaggaaagtggttttctataatgaataatgcccgcactaatgcaaaaagacgaagattatcttctaaacaagggggattaagcatatccgaaggaaaagagagtaatat
acccagtgttgttgaagaaagcaaggataatttggaacaagcttctgcagatgacaggctaaattttggtgaccgaattttggtaaaagccccaggttatccatggtggccg
gccttgctactgagacgaaaagaaactaaggatagtttgaatactaatagctcatttaatgtcttatataaggttttgtttttcctgacttcaattttgcatgggtgaaaagaaata
gtgttaagccattattggattccgaaatagccaaatttcttggttcctcaaagcggaagtctaaagaacttattgaagcttatgaggcttcaaaaactcctcctgatttaaaggag
gaatcttccaccgatgaggaaatggatagcttatcagctgctgaggagaagcctaattttttgcaaaaaagaaaatatcattgggagacatctcttgatgaatcagatgcgga
gagtatctccagcggatccttgatgtcaataacttctatttctgaaatgtatggtcctactgtcgcttcgacttctcgtagctctacgcagttaagtgaccaaaggtacc

Figure 53

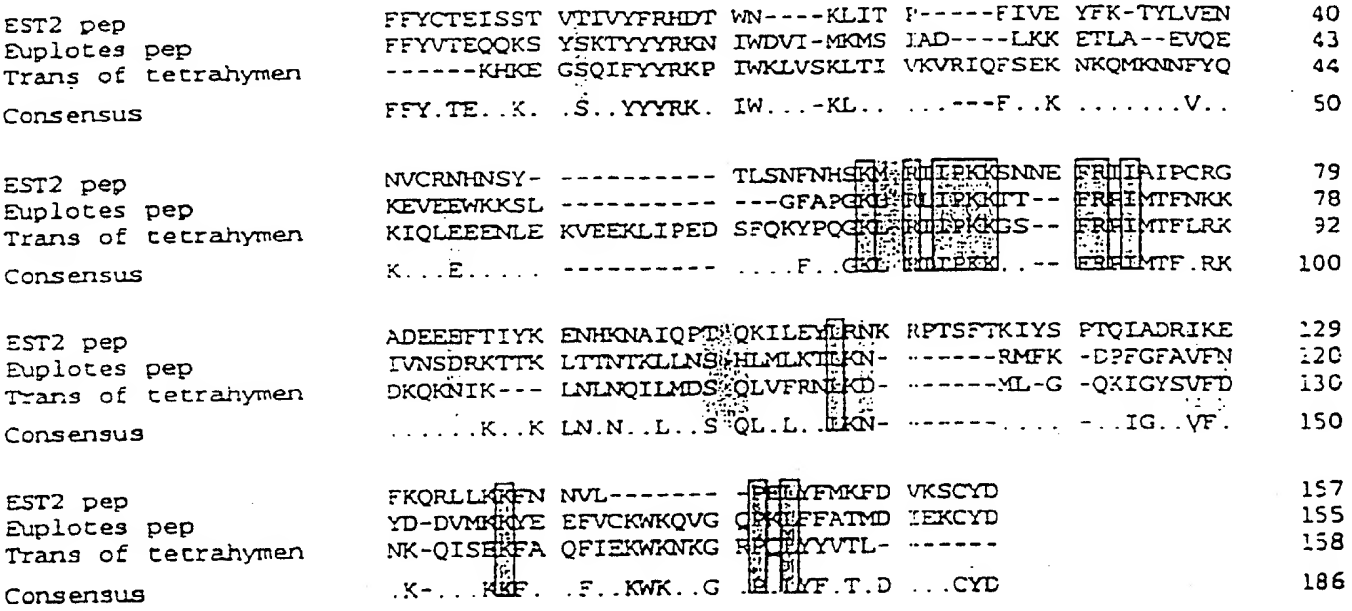

EST2 pep
Euplotes pep
Trans of tetrahymen
Consensus
EST2 pep
Euplotes pep
Trans of tetrahymen
Consensus
EST2 pep
Euplotes pep
Trans of tetrahymen
Consensus
EST2 pep
Euplotes pep
Trans of tetrahymen
Consensus

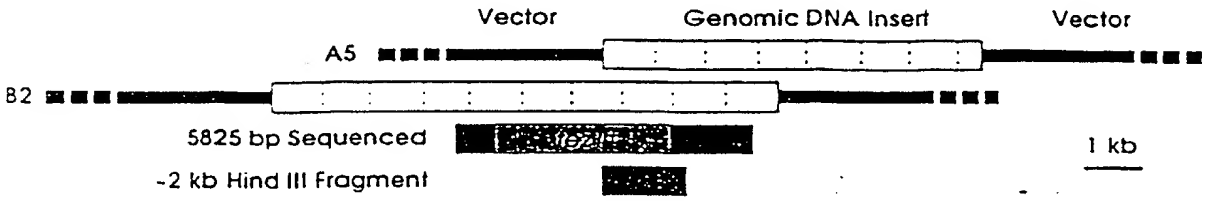
Figure 54

S-1: FFY VTE TTF QKN RLF FYR KSV WSK
S-2: RQH LKR VQL RDV SEA EVR QHR EA
S-3: AFT FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q
A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS
A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

Figure 55

A


Vector
Genomic DNA Insert
Vector
A5
B2
5825 bp Sequenced
tez1+
~2 kb Hind III Fragment
1 kb

B

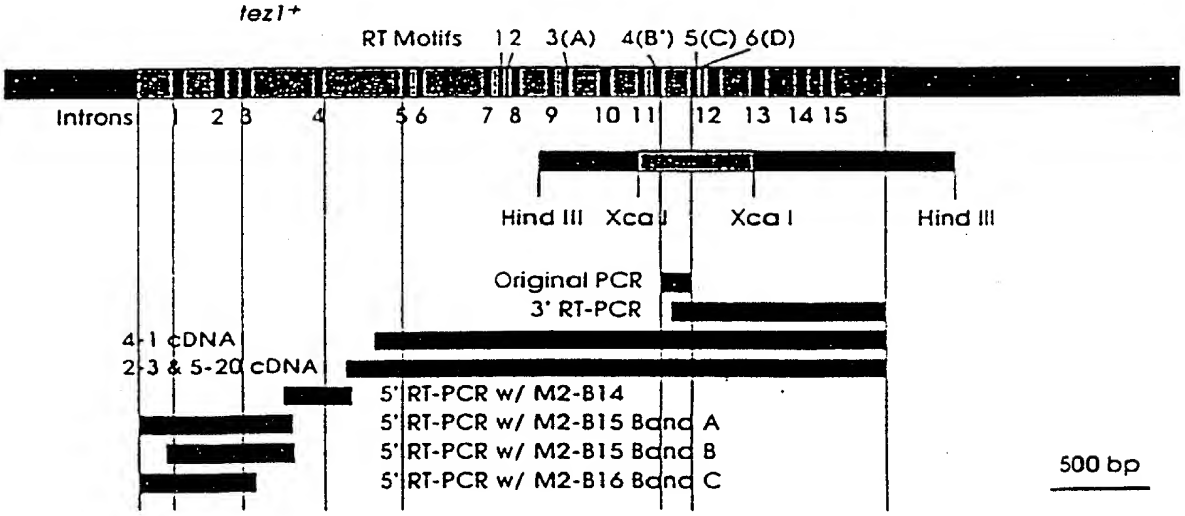

tez1+
RT Motifs 1 2 3(A) 4(B') 5(C) 6(D)
Introns 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Hind III
Xca I
Xca I
Hind III
Original PCR
3' RT-PCR
4-1 cDNA
2-3 & 5-20 cDNA
5' RT-PCR w/ M2-B14
5' RT-PCR w/ M2-B15 Band A
5' RT-PCR w/ M2-B15 Band B
5' RT-PCR w/ M2-B16 Band C
500 bp

Figure 56

Poly 4

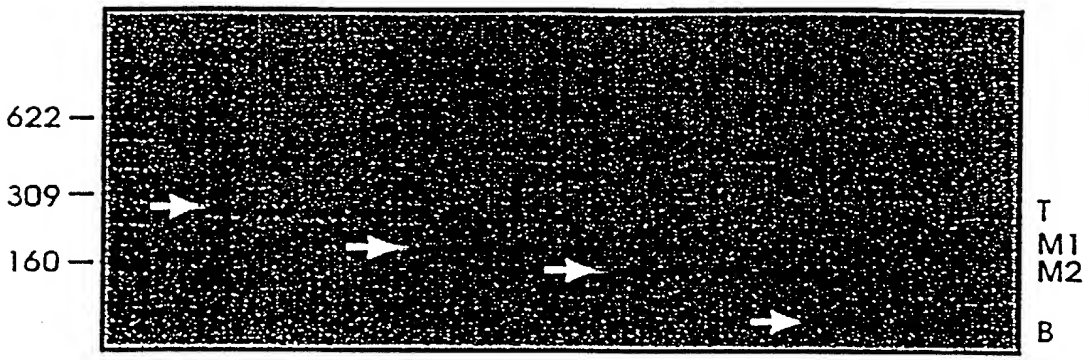
```
              t       t       c
    t a   a   g   c   c   t c g
5'- cag acc aaa gga att cca taa gg -3'
    Q   T   K   G   I   P   Q   G
```

4 (B')

5 (C')

```
    D   D   Y   L   L   I   T
3'- ctg ctg atg gag gag tag tgg -5'
      a   a   a a a a a   a   a
                  t   t   t   t
                          c   c
                        Poly 1
```

Figure 57



Motif B' (4)

QTKGIPQG

Motif C (5)

DDYLLIT

Figure 58

PCR Product M2 showed Reasonable Match with Other Telomerase Proteins

```
Ot                  LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
Ea_p123        KGIPQGLCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLIT
Sp_M2               SILSSFLCHFYMEDLIDEYLSFTKKK--------GSVLLRVV
Sc_p103        DGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPS------QDTLILKLADDFLIIS
                       .*   .        *  .     *                .....

Q   K   V   G   I   P   Q   G
caa aaa gtt ggt atc cct cag gg.....    <---Actual Genomic Sequence.

Poly 4
      t       t       c
t a   a   g   c   c   t c g
cag acc aaa gga att cca taa gg ---->

ag acc aaa gga att cca tca ggC TCA ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG
tc tgg ttt cct taa ggt agt ccG AGT TAA GAC AGT AGA AAA AAC ACA GTA AAG ATA TAC

      K   G   :   P   S   G   S   I   L   S   S   F   L   C   H   F   Y   M

GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC TTT CCT AGT CAC AAC AAT GCT

E   D   L   I   D   E   Y   L   S   F   T   K   K   K   G   S   V   L   L   R

GTA GTC gac gac tac ctc ctc atc acc
CAT CAG ctg ctg atg gag gag tag tgg

V   V   D   D   Y   L   L   I   T

  <---- ctg ctg atg gag gag tag tgg
        a   a   a a a a a   a   a
                t   t   t   t
                        c   c
                    Poly 1

   .....gac gat ttc ctc ttt ata aca.....  <---Actual Genomic Sequence.
        D   D   F   L   F   I   T
```

Figure 59

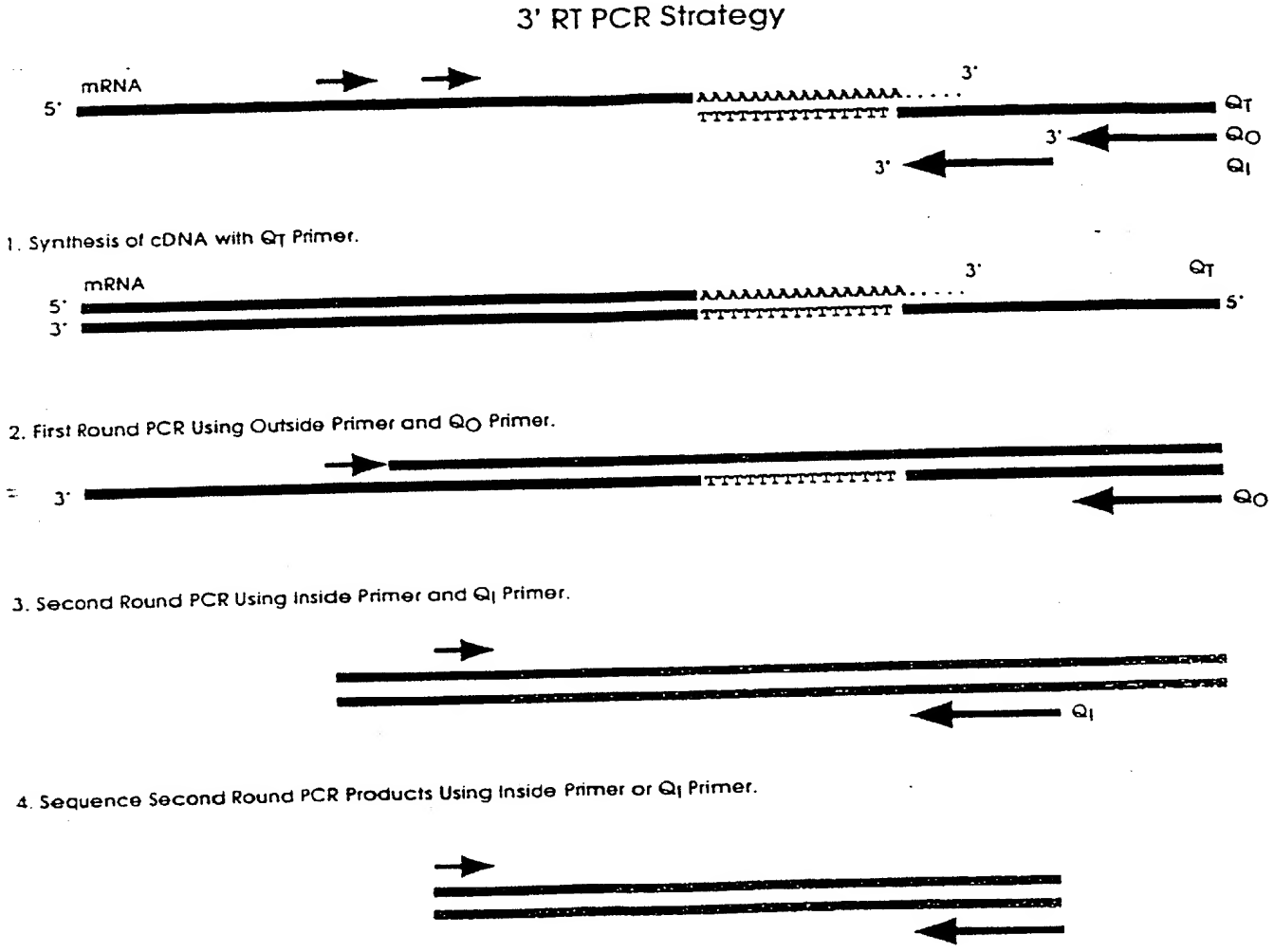

3' RT PCR Strategy
mRNA
5'
3'
AAAAAAAAAAAAAAAA.....
TTTTTTTTTTTTTTT
Q_T
3'
Q_O
3'
Q_I
1. Synthesis of cDNA with Q_T Primer.
mRNA
5'
3'
3'
AAAAAAAAAAAAAAAA.....
TTTTTTTTTTTTTTT
Q_T
5'
2. First Round PCR Using Outside Primer and Q_O Primer.
3'
TTTTTTTTTTTTTTT
Q_O
3. Second Round PCR Using Inside Primer and Q_I Primer.
Q_I
4. Sequence Second Round PCR Products Using Inside Primer or Q_I Primer.

Figure 60

A

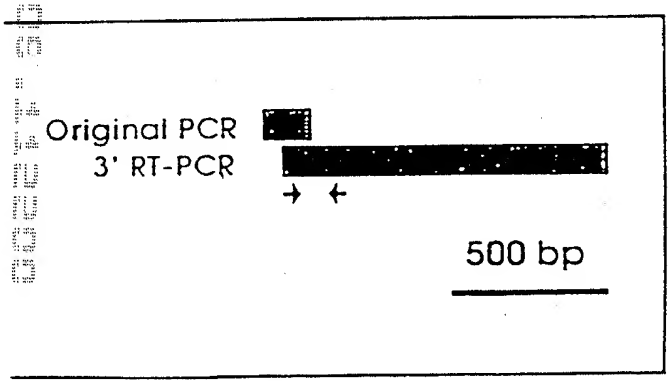
-Size Selected Libraries from P. Nurese
- 3 ~ 4 kb
- 5 ~ 6 kb
- 7 ~ 8 kb
- 11 ~ 12 kb

-Libraries from J.A. Wise
- Sau 3a Partial Digest
- Hind III Partial Digest

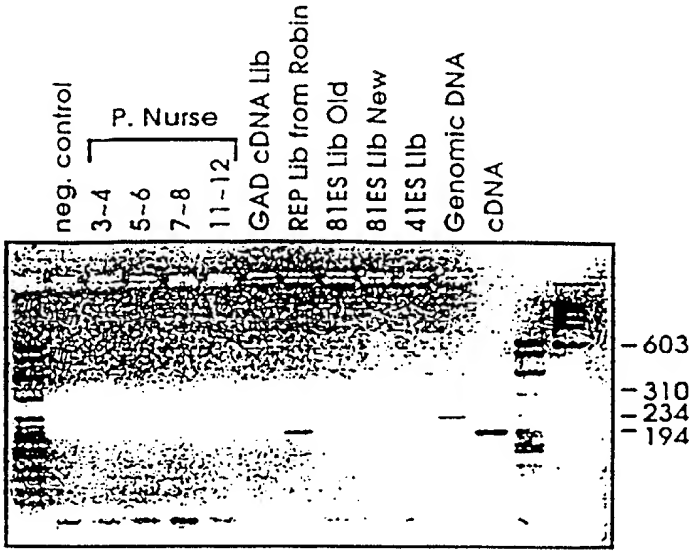
cDNA Libraries

- GAD (Gal Activation Domain) Library
- REP Library from R. Allshire
- REP81ES Library (old)
- REP81ES Library (new)
- REP41ES Library

B



C



D

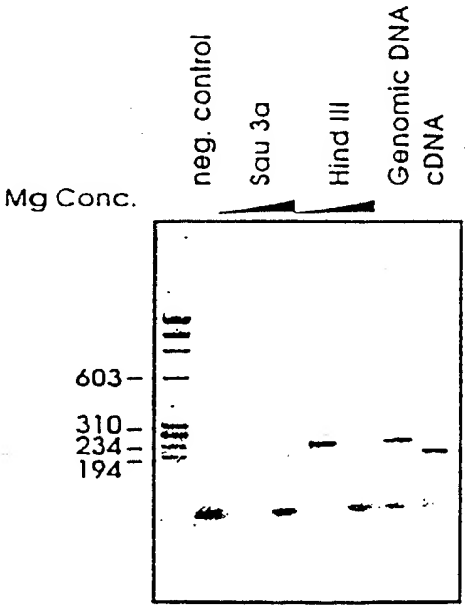
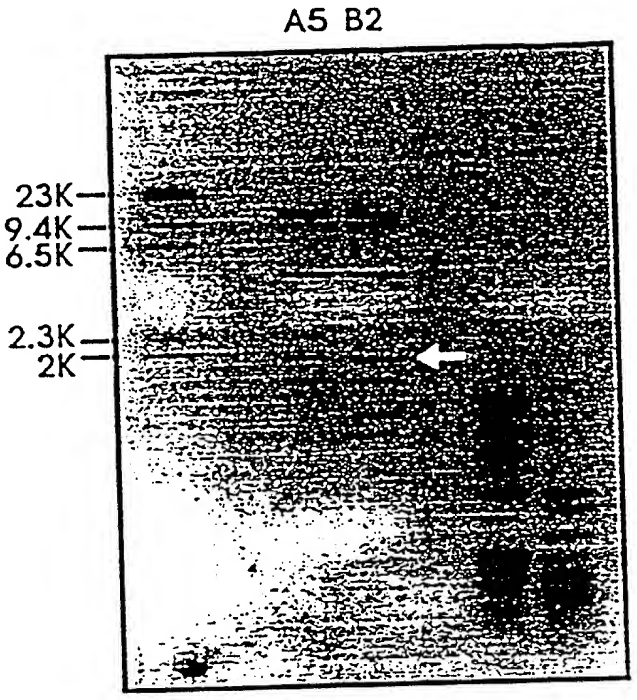


Figure 61



Hind III Digested Positive Genomic Clones

Figure 62

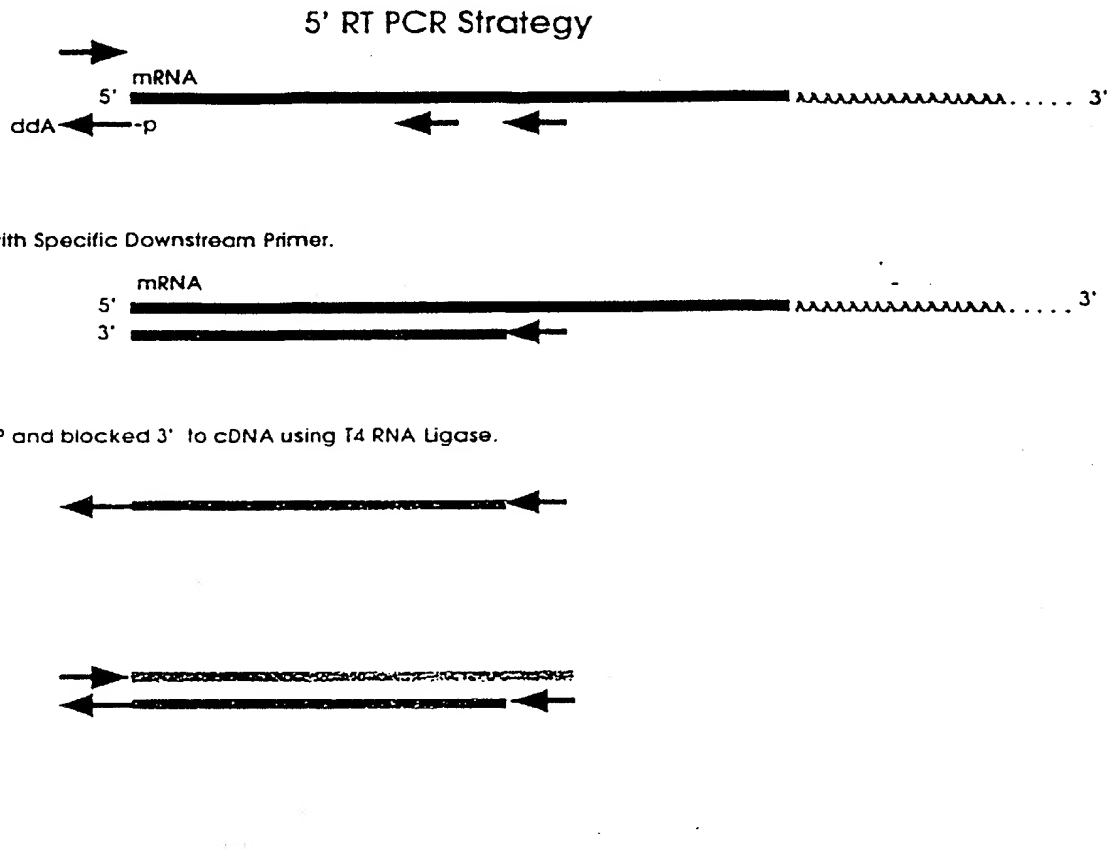

5' RT PCR Strategy
mRNA
5'
3'
ddA
-p
1. Synthesis of cDNA with Specific Downstream Primer.
mRNA
5'
3'
2. Ligate Oligo with 5'-P and blocked 3' to cDNA using T4 RNA Ligase.
3. First Round PCR
4. Second Round PCR

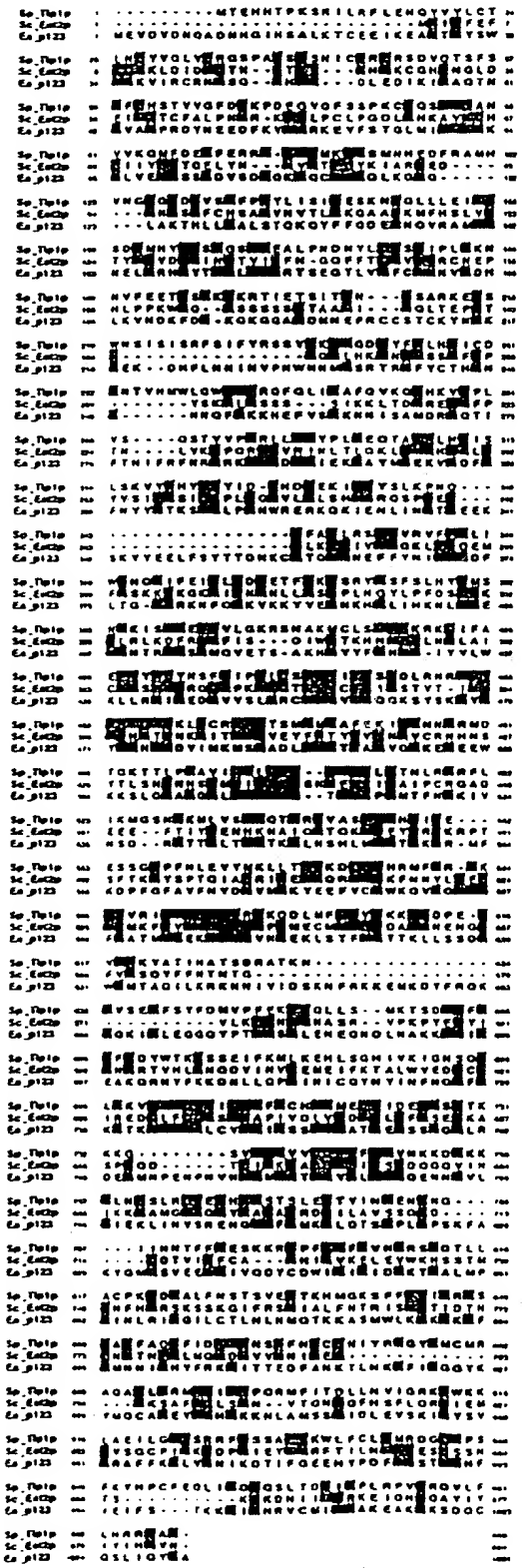
Figure 63

Alignment of RT Domains from Telomerase Catalytic Subunits.

```
                                    Motif O
S.p. Tez1p (429). WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIW ...(35)...
S.c. Est2p (366). WLFRQLIPKIIQTFFYCTEISSTVT-IVYFRHDTW ...(35)...
E.a. p123  (441). WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIW ...(35)...
                  *                ***  **          * *   *

                   Motif 1     Motif 2      K
                  p hh h K      hR h        R
S.p. Tez1p  AVIRLLPKK--NTFRLITN-LRKRF ...(61)...
S.c. Est2p  SKMRIIPKKSNNEFRIIAIPCRGAD ...(62)...
E.a. p123   GKLRLIPKK--TTFRPIMTFNKKIV ...(61)...
               *  ***    ** *
               Motif 3(A) AF
                  h  hDh  GY  h
S.p. Tez1p  KKYFVRIDIKSCYDRIKQDLMFRIVK ...(89)...
S.c. Est2p  ELYFMKFDVKSCYDSIPRMECMRILK ...(75)...
E.a. p123   KLFFATMDIEKCYDSVNREKLSTFLK ...(107)...
               *   *   ***           *
                Motif 4(B')
                    hPQG   pP hh   h
S.p. Tez1p  YLQKVGIPQGSILSSFLCHFYMEDLIDEYLSF ...(6)...
S.c. Est2p  YIREDGLFQGSSLSAPIVDLVYDDLLEFYSEF ...(8)...
E.a. p123   YKQTKGIPQGLCVSSILSSFYYATLEESSLGF ...(14)...
            *   *  **   *          *       *
                  Y Motif 5(C)                    Motif 6(D)
                 h  F DDhhh                      Gh h  cK  h
S.p. Tez1p  VLLRVVDDFLFITVNKKDAKKFLNLSLRGFEKHNFSTSLEKTVINFENS .(205)
S.c. Est2p  LILKLADDFLIISTDQQQVINIKKLAMGGFQKYNAKANRDKILAVSSQS .(173)
E.a. p123   LLMRLTDDYLLITTQENNAVLFIEKLINVSRENGFKFNMKKLQTSFPLS .(209)
               ** * *                                  *       *
```

Figure 64

A



B

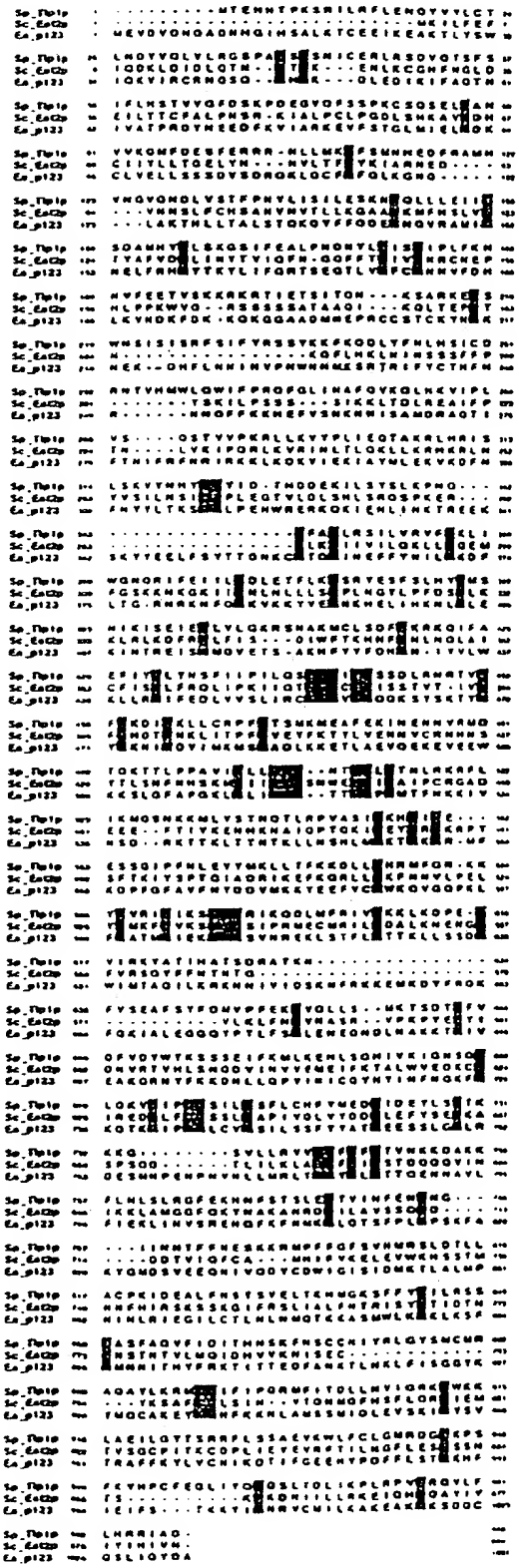
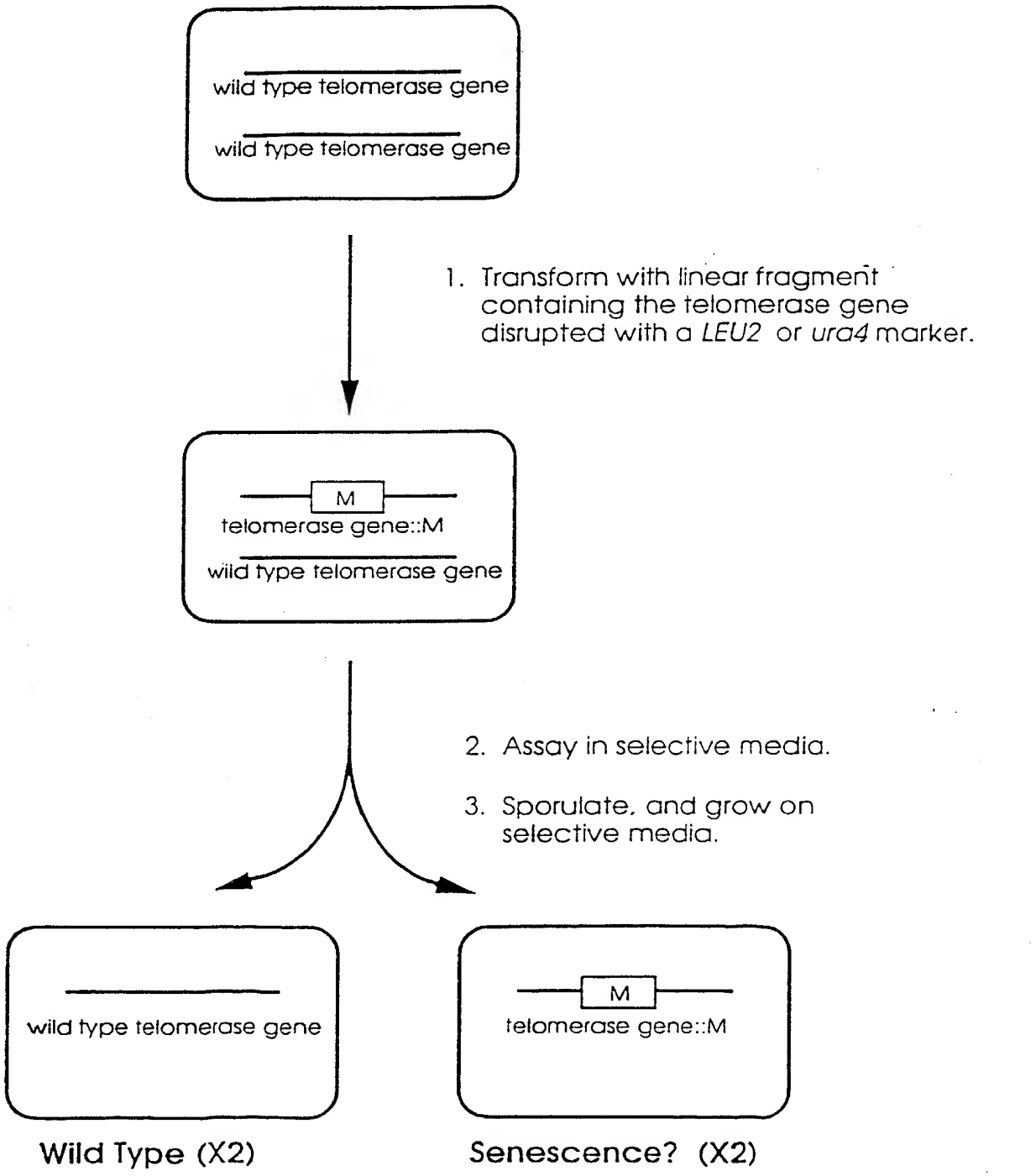


Figure 65

Disruption strategy for the putative telomerase genes.



(These cells will show a senescence phenotype if the disrupted gene encodes a telomerase subunit.)

Figure 66

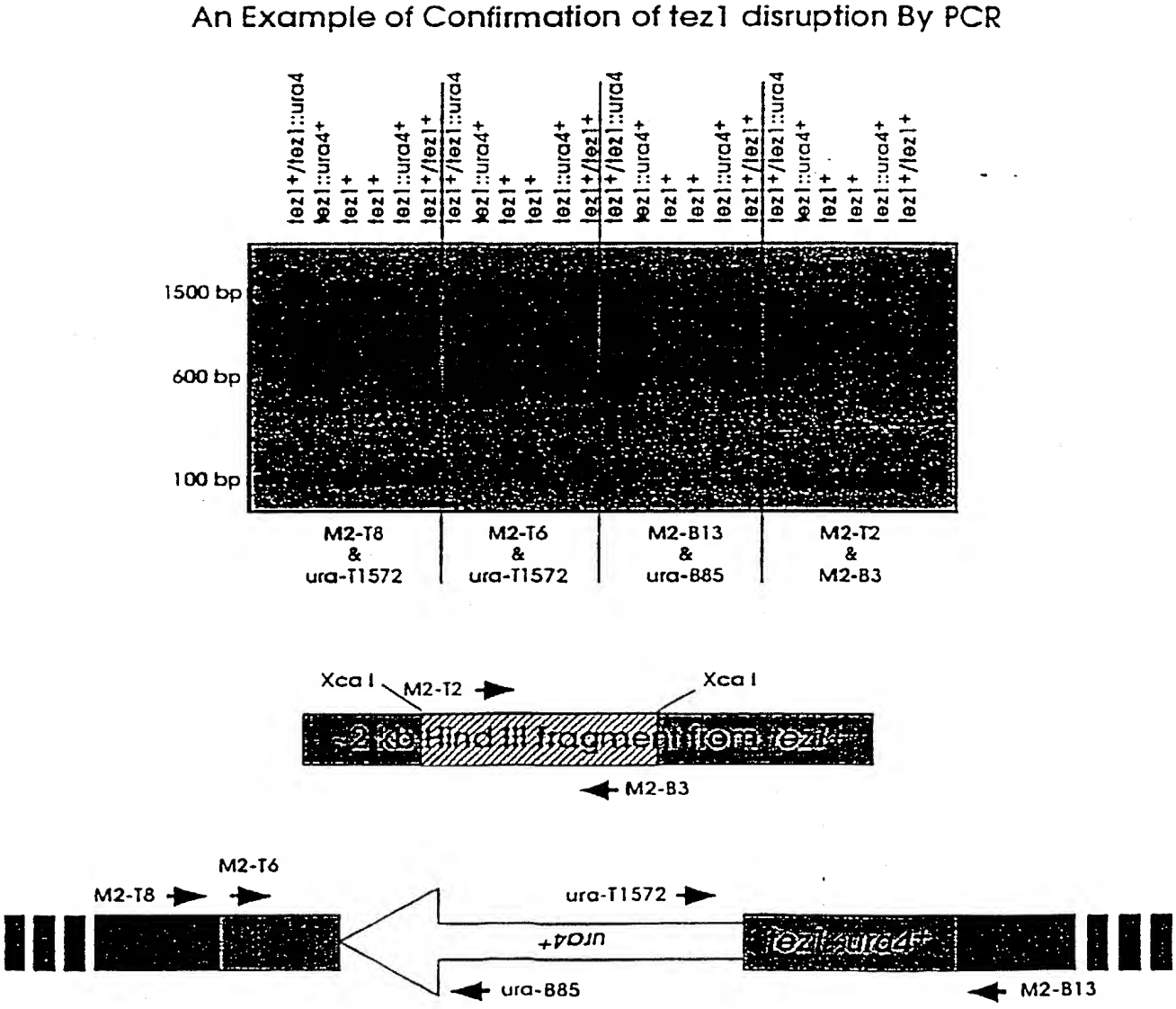

An Example of Confirmation of tez1 disruption By PCR
tez1+/tez1::ura4
tez1::ura4+
tez1+
1500 bp
600 bp
100 bp
M2-T8 & ura-T1572
M2-T6 & ura-T1572
M2-B13 & ura-B85
M2-T2 & M2-B3
Xca I
M2-T2
Xca I
~2 kb HindIII fragment from tez1+
M2-B3
M2-T6
M2-T8
ura-T1572
ura4+
tez1::ura4+
ura-B85
M2-B13

Figure 67

Tez1 disruption causes progressive shortening of telomeres in *S. pombe*

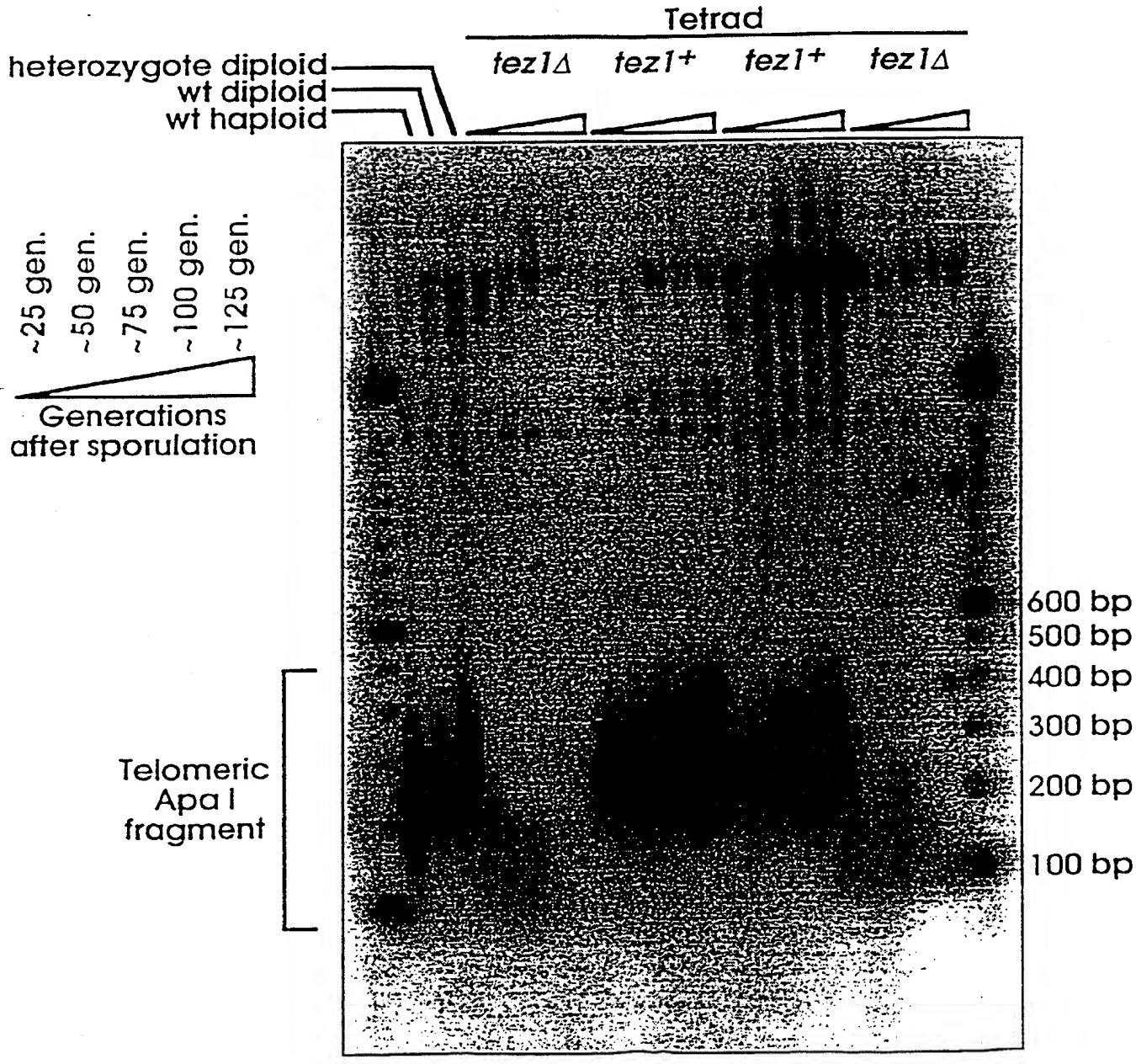


Figure 68

```
                                1
                               met ser val tyr val val glu leu leu
GCCAAGTTCCTGCACTGGCTG          ATG AGT GTG TAC GTC GTC GAG CTG CTC

 10                                                  20
arg ser phe phe tyr val thr glu thr thr phe gln lys asn arg
AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG

                        30
leu phe phe tyr arg lys ser val trp ser lys leu gln ser ile
CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT

 40                                                  50
gly ile arg gln his leu lys arg val gln leu arg glu leu ser
GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG TCG

                        60
glu ala glu val arg gln his arg glu ala arg pro ala leu leu
GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG CTG

 70                                                  80
thr ser arg leu arg phe ile pro lys pro asp gly leu arg pro
ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG CCG

                        90
ile val asn met asp tyr val val gly ala arg thr phe arg arg
ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC AGA

100                                                 110
glu lys     ala glu arg leu thr ser arg val lys ala leu phe
GAA AAG ARG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG TTC

                       120
ser val leu asn tyr glu arg ala arg arg pro gly leu leu gly
AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG GGC

130                                                 140
ala ser val leu gly leu asp asp ile his arg ala trp arg thr
GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC ACC

                       150
phe val leu arg val arg ala gln asp pro pro pro glu leu tyr
TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG TAC

160                                                 170
phe val lys val asp val thr gly ala tyr asp thr ile pro gln
TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG

                       180
asp arg leu thr glu val ile ala ser ile ile lys pro gln asn
GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG AAC
```

Figure 68 (cont.)

```
190                                                 200
thr tyr cys val arg arg tyr ala val val gln lys ala ala met
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG

                        210
gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys
GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC

220                                                 230
gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser
CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC

                        240
leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg
CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG

250                                                 260
arg asp gly leu leu leu arg leu val asp asp phe leu leu val
CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG TTG GTG

                        270
thr pro his leu thr his ala lys thr phe leu arg thr leu val
ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC

280                                                 290
arg gly val pro glu tyr gly cys val val asn leu arg lys thr
CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG AAG ACA

                        300
val val asn phe pro val glu asp glu ala leu gly gly thr ala
GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT

310                                                 320
phe val gln met pro ala his gly leu phe pro trp cys gly leu
TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG

                        330
leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser
CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC

340                                                 350
tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly
TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC

                        360
phe lys ala gly arg asn met arg arg lys leu phe gly val leu
TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG

370                                                 380
arg leu lys cys his ser leu phe leu asp leu gln val asn ser
CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC
```

Figure 68 (cont.)

```
                        390
leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln
CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

400                                             410
ala tyr arg phe his ala cys val leu gln leu pro phe his gln
GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

                        420
gln val trp lys asn pro his phe ser cys ala ser ser leu thr
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430                                             440
arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

                        450
val ala gly gly gln gly arg arg arg pro ser ala leu arg gly
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460                                             470
arg ala val ala val pro pro ser ile pro ala gln ala asp ser
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

                        480
thr pro cys his leu arg ala thr pro gly val thr gln asp ser
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490                                             500
pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

                        510
pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520                                             530
his pro gly leu met ala thr arg pro gln pro gly arg glu gln
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

                        540
thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550                                             560
arg gly gly pro his pro gly leu his arg trp glu ser glu ala
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

564
OP
TGA GTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC

CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTTCACTTCCCCAC
```

Figure 68 (cont.)

AGGCTGGCGTTCGGTCCACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACT

CCCCACATAGGAATAGTCCATCCCCAGATTCGCCATTGTTCACCCTTCGCCCTGCCTTCC

TTTGCCTTCCACCCCCACCATTCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTTTGGG

AATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGG

GTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATGAGTT

TTTCAGTTTTGGAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 69

```
Motif -1
Ep p123      ...LVVSLIRCFFYVTEQQKSYSKT...
Sp Tez1      ...FIIPILQSFFYITESSDLRNRT...
Sc Est2      ...LIPKIIQTFFYCTEISSTVTIV...
Hs TCP1      ...YVVELLRSFFYVTETTFQKNRL...
consensus        FFY TE

                        K
Motif 0        p hhh  K    hR h      R
Ep p123      ...KSLGFAPGKLRLIPKKT--TFRPIMTFNKKIV...
Sp Tez1      ...QKTTLPPAVIRLLPKKN--TFRLITNLRKRFL...
Sc Est2      ...TLSNFNHSKMRIIPKKSNNEFRIIAIPCRGAD...
Hs TCP1      ...ARPALLTSRLRFIPKPD--GLRPIVNMDYVVG...
consensus          R PK    R I

                   AF
Motif A          h  hDh  GY  h
Ep p123      ...PKLFFATMDIEKCYDSVNREKLSTFLK...
Sp Tez1      ...RKKYFVRIDIKSCYDRIKQDLMFRIVK...
Sc Est2      ...PELYFMKFDVKSCYDSIPRMECMRILK...
Hs TCP1      ...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...
consensus      F  D   YD

Motif B            hPQG   pS hh
Ep p123      ...NGKFYKQTKGIPQGLCVSSILSSFYYA...
Sp Tez1      ...GNSQYLQKVGIPQGSILSSFLCHFYME...
Sc Est2      ...EDKCYIREDGLFQGSSLSAPIVDLVYD...
Hs TCP1      ...RATSYVQCQGIPQGSILSTLLCSLCYG...
consensus          G  QG  S

                Y
Motif C          h  F DDhhh
Ep p123      ...PNVNLLMRLTDDYLLITTQENN...

Sp Tez1      ...KKGSVLLRVVDDFLFITVNKKD...
Sc Est2      ...SQDTLILKLADDFLIISTDQQQ...
Hs TCP1      ...RRDGLLLRLVDDFLLVTPHLTH...
consensus          DD L

Motif D          Gh h  cK
Ep p123      ...NVSRENGFKFNMKKL...
Sp Tez1      ...LNLSLRGFEKHNFST...
Sc Est2      ...KKLAMGGFQKYNAKA...
Hs TCP1      ...LRTLVRGVPEYGCVV...
consensus        G
```

Figure 70

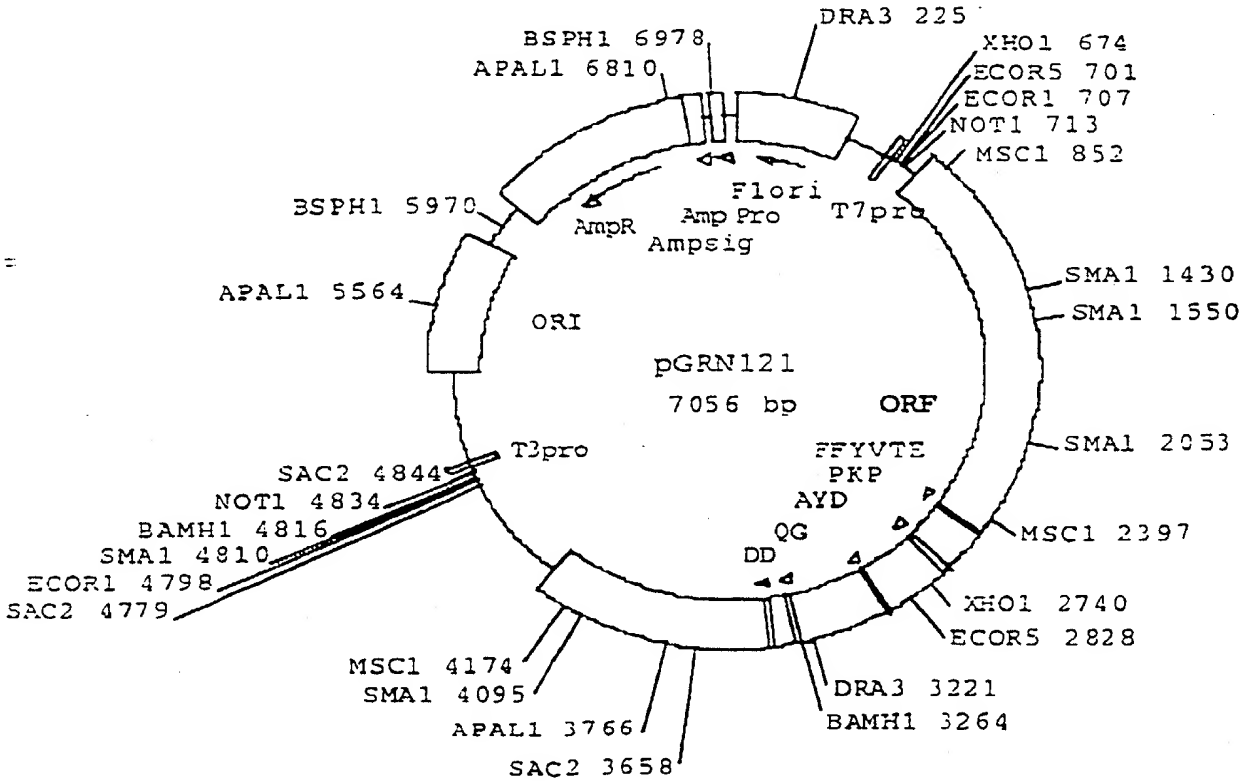

DRA3 225
BSPH1 6978
APAL1 6810
XHO1 674
ECOR5 701
ECOR1 707
NOT1 713
MSC1 852
Flori
BSPH1 5970
AmpR
Amp Pro
T7pro
Ampsig
SMA1 1430
SMA1 1550
APAL1 5564
ORI
pGRN121
7056 bp
ORF
SMA1 2053
T3pro
FFYVTE
PKP
SAC2 4844
NOT1 4834
AYD
MSC1 2397
BAMH1 4816
QG
SMA1 4810
DD
ECOR1 4798
XHO1 2740
SAC2 4779
ECOR5 2828
MSC1 4174
DRA3 3221
SMA1 4095
BAMH1 3264
APAL1 3766
SAC2 3658

Figure 71

```
   1 GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC
  51 CCGCGATGCC GCGCGCTCCC CGCTGCCGAG CCGTGCGCTC CCTGCTGCGC
 101 AGCCACTACC GCGAGGTGCT GCCGCTGGCC ACGTTCGTGC GGCGCCTGGG
 151 GCCCCAGGGC TGGCGGCTGG TGCAGCGCGG GGACCCGGCG GCTTTCCGCG
 201 CGNTGGTGGC CCANTGCNTG GTGTGCGTGC CCTGGGANGN ANGGCNGCCC
 251 CCCGCCGCCC CCTCCTTCCG CCAGGTGTCC TGCCTGAANG ANCTGGTGGC
 301 CCGAGTGCTG CANANGCTGT GCGANCGCGG CGCGAANAAC GTGCTGGCCT
 351 TCGGCTTCGC GCTGCTGGAC GGGGCCCGCG GGGGCCCCCC CGAGGCCTTC
 401 ACCACCAGCG TGCGCAGCTA CCTGCCCAAC ACGGTGACCG ACGCACTGCG
 451 GGGGAGCGGG GCGTGGGGGC TGCTGCTGCG CCGCGTGGGC GACGACGTGC
 501 TGGTTCACCT GCTGGCACGC TGCGCGNTNT TTGTGCTGGT GGNTCCCAGC
 551 TGCGCCTACC ANGTGTGCGG GCCGCCGCTG TACCAGCTCG GCGCTGCNAC
 601 TCAGGCCCGG CCCCCGCCAC ACGCTANTGG ACCCGAANGC GTCTGGGATC
 651 CAACGGGCCT GGAACCATAG CGTCAGGGAG GCCGGGGTCC CCCTGGGCTG
 701 CCAGCCCCGG GTGCGAGGAG GCGCGGGGGC AGTGCCAGCC GAAGTCTGCC
 751 GTTGCCCAAG AGGCCCAGGC GTGGCGCTGC CCCTGAGCCG GAGCGGACGC
 801 CCGTTGGGCA GGGGTCCTGG GCCCACCCGG GCAGGACGCC TGGACCGAGT
 851 GACCGTGGTT TCTGTGTGGT GTCACCTGCC AGACCCGCCG AAGAAGCCAC
 901 CTCTTTGGAG GGTGCGCTCT CTGGCACGCG CCACTCCCAC CCATCCGTGG
 951 GCCGCCAGCA CCACGCGGGC CCCCCATCCA CATCGCGGCC ACCACGTCCT
1001 GGGACACGCC TTGTCCCCCG GTGTACGCCG AGACCAAGCA CTTCCTCTAC
1051 TCCTCAGGCG ACAAGNACAC TGCGNCCCTC CTTCCTACTC AATATATCTG
1101 AGGCCCAGCC TGACTGGCGT TCGGGAGGTT CGTGGAGACA NTCTTTCTGG
1151 TTCCAGGCCT TGGATGCCAG GATTCCCCGC AGGTTGCCCC GCCTGCCCCA
1201 GCGNTACTGG CAAATGCGGC CCCTGTTTCT GGAGCTGCTT GGGAACCACG
1251 CGCAGTGCCC CTACGGGGTG TTCCTCAAGA CGCACTGCCC GCTGCGAGCT
1301 GCGGTCACCC CAGCAGCCGG TGTCTGTGCC CGGGAGAAGC CCCAGGGCTC
1351 TGTGGCGGCC CCCGAGGAGG AGGAACACAG ACCCCCGTCG CCTGGTGCAG
1401 CTGCTCCGCC AGCACAGCAG CCCCTGGCAG GTGTACGGCT TCGTGCGGGC
1451 CTGCCTGCGC CGGCTGGTGC CCCCAGGCCT CTGGGGCTCC AGGCACAACG
1501 AACGCCGCTT CCTCAGGAAC ACCAAGAAGT TCATCTCCCT GGGGAAGCAT
1551 GCCAAGCTCT CGCTGCAGGA GCTGACGTGG AAGATGAGCG TGCGGGACTG
1601 CGCTTGGCTG CGCAGGAGCC CAGGGGTTGG CTGTGTTCCG GCCGCAGAGC
1651 ACCGTCTGCG TGAGGAGATC CTGGCCAAGT TCCTGCACTG GCTGATGAGT
1701 GTGTACGTCG TCGAGCTGCT CAGGTCTTTC TTTTATGTCA CGGAGACCAC
1751 GTTTCAAAAG AACAGGCTCT TTTTCTACCG GAAGAGTGTC TGGAGCAAGT
1801 TGCAAAGCAT TGGAATCAGA CAGCACTTGA AGAGGGTGCA GCTGCGGGAG
1851 CTGTCGGAAG CAGAGGTCAG GCAGCATCGG GAAGCCAGGC CCGCCCTGCT
1901 GACGTCCAGA CTCCGCTTCA TCCCCAAGCC TGACGGGCTG CGGCCGATTG
1951 TGAACATGGA CTACGTCGTG GGAGCCAGAA CGTTCCGCAG AGAAAAGAGG
2001 GCCGAGCGTC TCACCTCGAG GGTGAAGGCA CTGTTCAGCG TGCTCAACTA
2051 CGAGCGGGCG CGGCGCCCCG GCCTCCTGGG CGCCTCTGTG CTGGGCCTGG
2101 ACGATATCCA CAGGGCCTGG CGCACCTTCG TGCTGCGTGT GCGGGCCCAG
2151 GACCCGCCGC CTGAGCTGTA CTTTGTCAAG GTGGATGTGA CGGGCGCGTA
2201 CGACACCATC CCCCAGGACA GGCTCACGGA GGTCATCGCC AGCATCATCA
2251 AACCCCAGAA CACGTACTGC GTGCGTCGGT ATGCCGTGGT CCAGAAGGCC
```

Figure 71 (cont.)

2301 GCCCATGGGC ACGTCCGCAA GGCCTTCAAG AGCCACGTCT CTACCTTGAC
2351 AGACCTCCAG CCGTACATGC GACAGTTCGT GGCTCACCTG CAGGANAACA
2401 GCCCGCTGAG GGATGCCGTC GTCATCGAGC AGAGCTCCTC CCTGAATGAG
2451 GCCAGCAGTG GCCTCTTCGA CGTCTTCCTA CGCTTCATGT GCCACCACGC
2501 CGTGCGCATC AGGGGCAAGT CCTACGTCCA GTGCCAGGGG ATCCCGCAGG
2551 GCTCCATCCT CTCCACGCTG CTCTGCAGCC TGTGCTACGG CGACATGGAG
2601 AACAAGCTGT TTGCGGGGAT TCGGCGGGAC GGGCTGCTCC TGCGTTTGGT
2651 GGATGATTTC TTGTTGGTGA CACCTCACCT CACCCACGCG AAAACCTTCC
2701 TCAGGACCCT GGTCCGAGGT GTCCCTGAGT ATGGCTGCGT GGTGAACTTG
2751 CGGAAGACAG TGGTGAACTT CCCTGTAGAA GACGAGGCCC TGGGTGGCAC
2801 GGCTTTTGTT CAGATGCCGG CCCACGGCCT ATTCCCCTGG TGCGGCCTGC
2851 TGCTGGATAC CCGGACCCTG GAGGTGCAGA GCGACTACTC CAGCTATGCC
2901 CGGACCTCCA TCAGAGCCAG TCTCACCTTC AACCGCGGCT TCAAGGCTGG
2951 GAGGAACATG CGTCGCAAAC TCTTTGGGGT CTTGCGGCTG AAGTGTCACA
3001 GCCTGTTTCT GGATTTGCAG GTGAACAGCC TCCAGACGGT GTGCACCAAC
3051 ATCTACAAGA TCCTCCTGCT GCAGGCGTAC AGGTTTCACG CATGTGTGCT
3101 GCAGCTCCCA TTTCATCAGC AAGTTTGGAA GAACCCCACA TTTTTCCTGC
3151 GCGTCATCTC TGACACGGCC TCCCTCTGCT ACTCCATCCT GAAAGCCAAG
3201 AACGCAGGGA TGTCGCTGGG GGCCAAGGGC GCCGCCGGCC CTCTGCCCTC
3251 CGAGGCCGTG CAGTGGCTGT GCCACCAAGC ATTCCTGCTC AAGCTGACTC
3301 GACACCGTGT CACCTACGTG CCACTCCTGG GGTCACTCAG GACAGCCCAG
3351 ACGCAGCTGA GTCGGAAGCT CCCGGGGACG ACGCTGACTG CCCTGGAGGC
3401 CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT
3451 GGCCACCCGC CCACAGCCAG GCCGAGAGCA GACACCAGCA GCCCTGTCAC
3501 GCCGGGCTCT ACGTCCCAGG GAGGGAGGGG CGGCCCACAC CCAGGCCCGC
3551 ACCGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCGAG GCCTGCATGT
3601 CCGGCTGAAG GCTGAGTGTC CGGCTGAGGC CTGAGCGAGT GTCCAGCCAA
3651 GGGCTGAGTG TCCAGCACAC CTGCCGTCTT CACTTCCCCA CAGGCTGGCG
3701 CTCGGCTCCA CCCCAGGGCC AGCTTTTCCT CACCAGGAGC CCGGCTTCCA
3751 CTCCCCACAT AGGAATAGTC CATCCCCAGA TTCGCCATTG TTCACCCCTC
3801 GCCCTGCCCT CCTTTGCCTT CCACCCCCAC CATCCAGGTG GAGACCCTGA
3851 GAAGGACCCT GGGAGCTCTG GGAATTTGGA GTGACCAAAG GTGTGCCCTG
3901 TACACAGGCG AGGACCCTGC ACCTGGATGG GGGTCCCTGT GGGTCAAATT
3951 GGGGGGAGGT GCTGTGGGAG TAAAATACTG AATATATGAG TTTTTCAGTT
4001 TTGAAAAAAA AAAAAAAAAA AAAAAAAAA

Figure 72

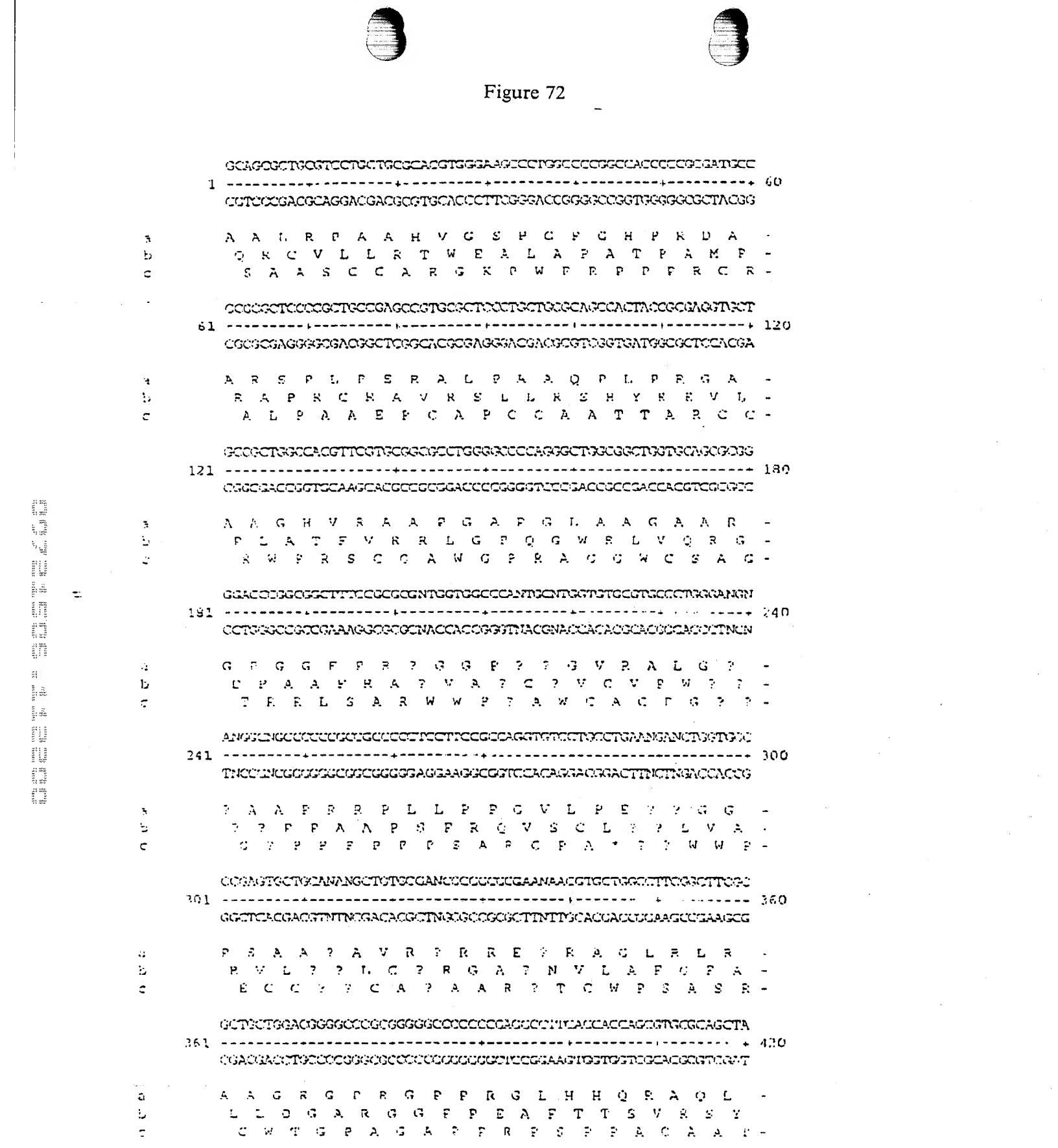


Figure 72 (cont.)

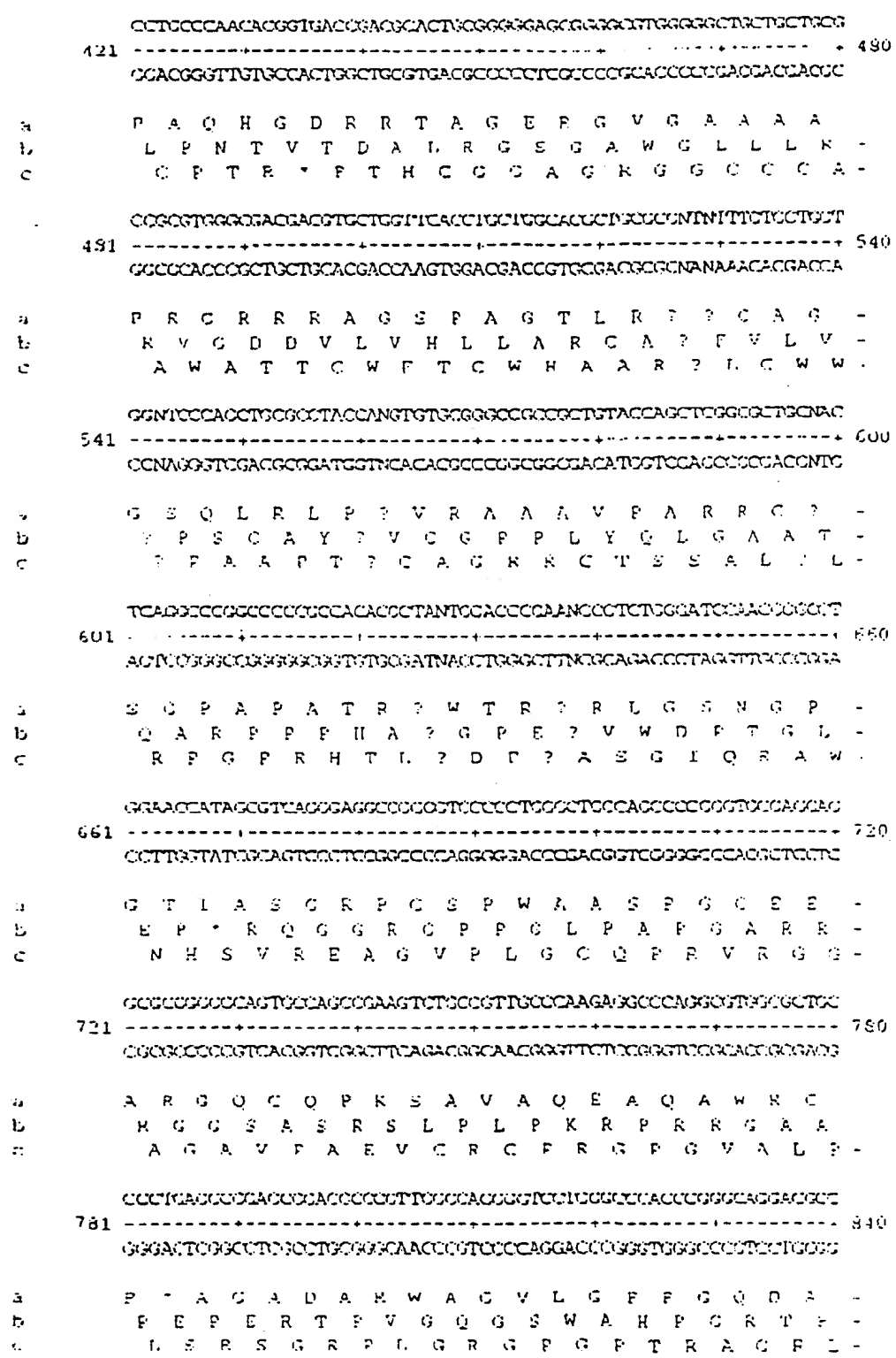


Figure 72 (cont.)

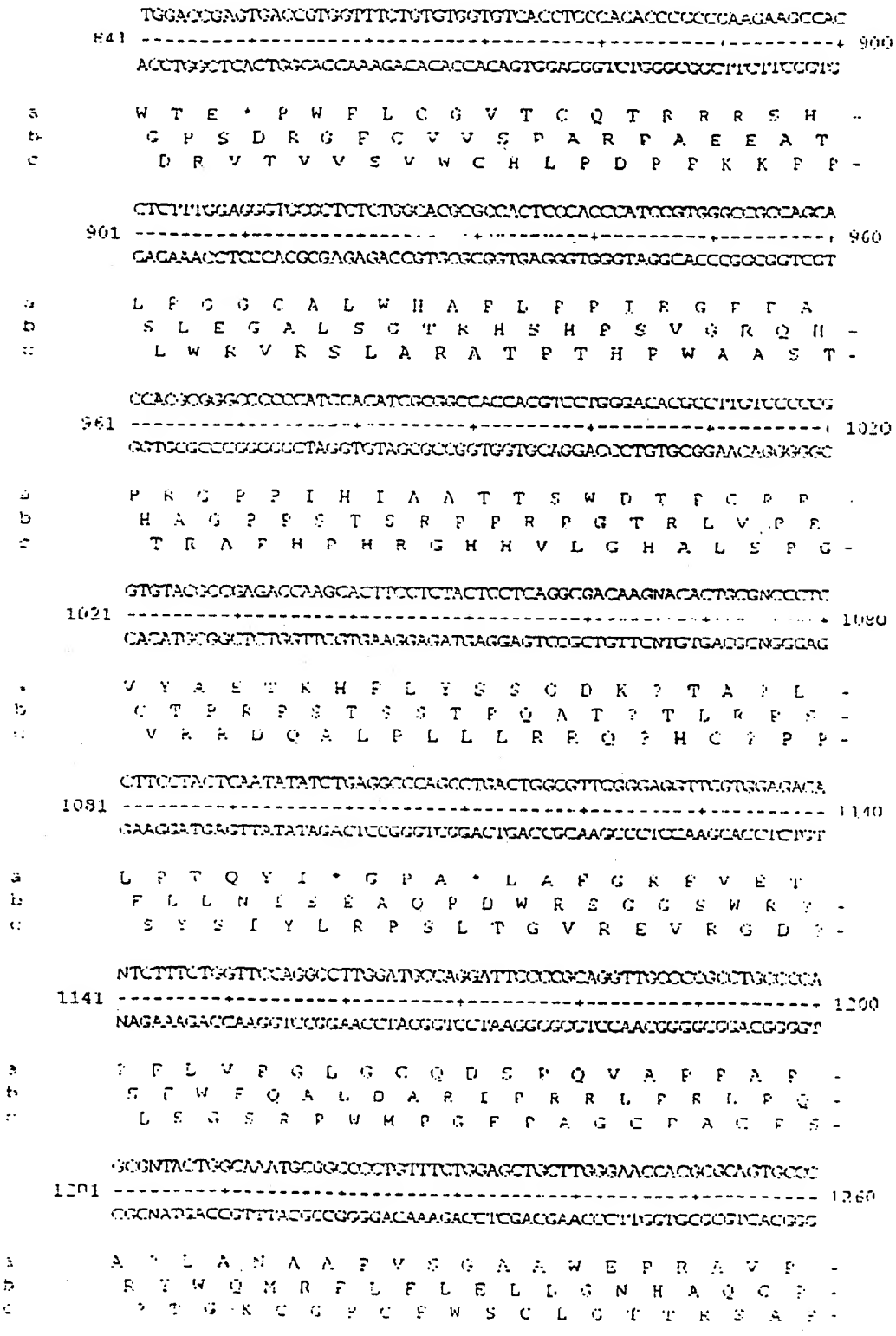


Figure 72 (cont.)

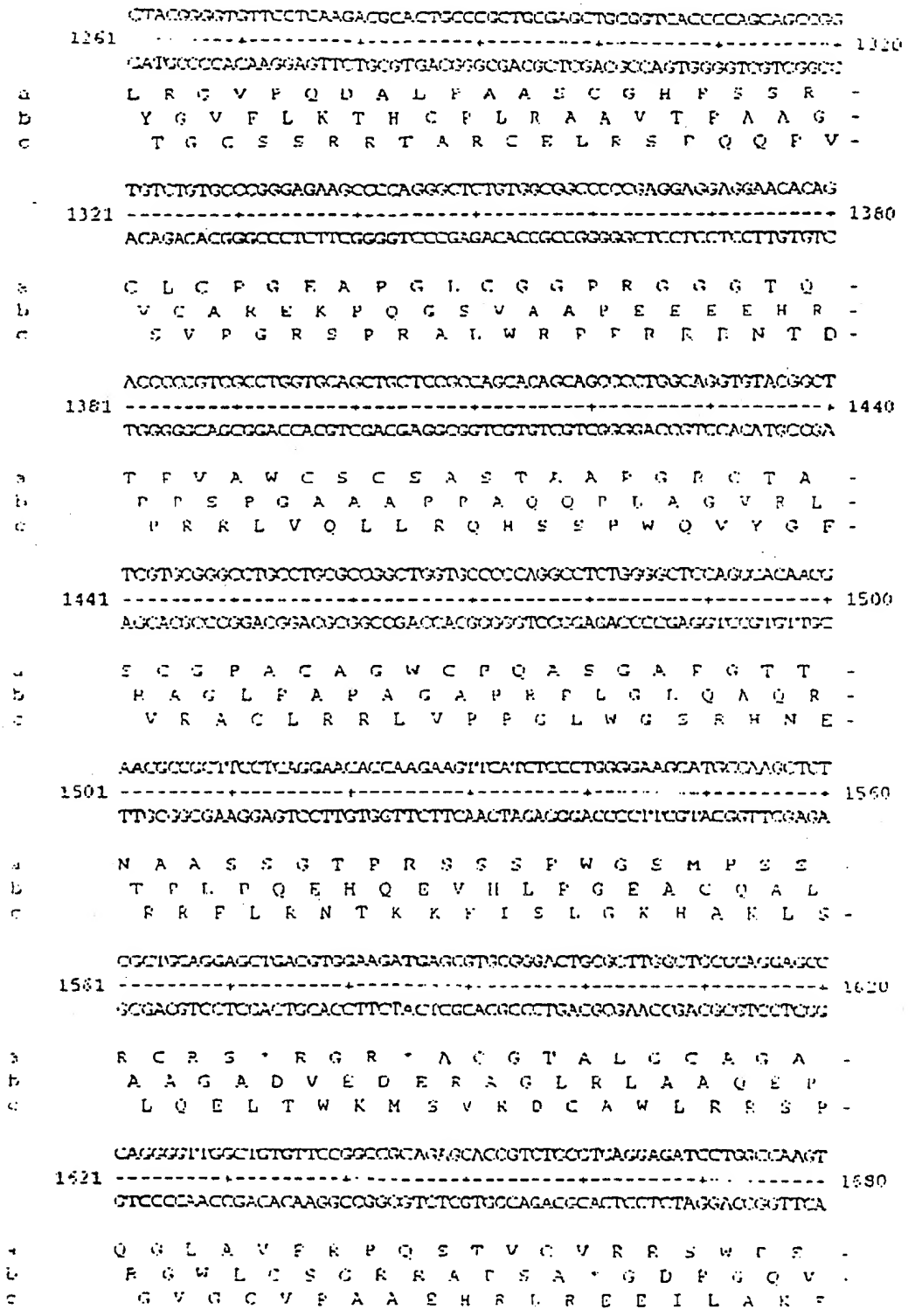


Figure 72 (cont.)

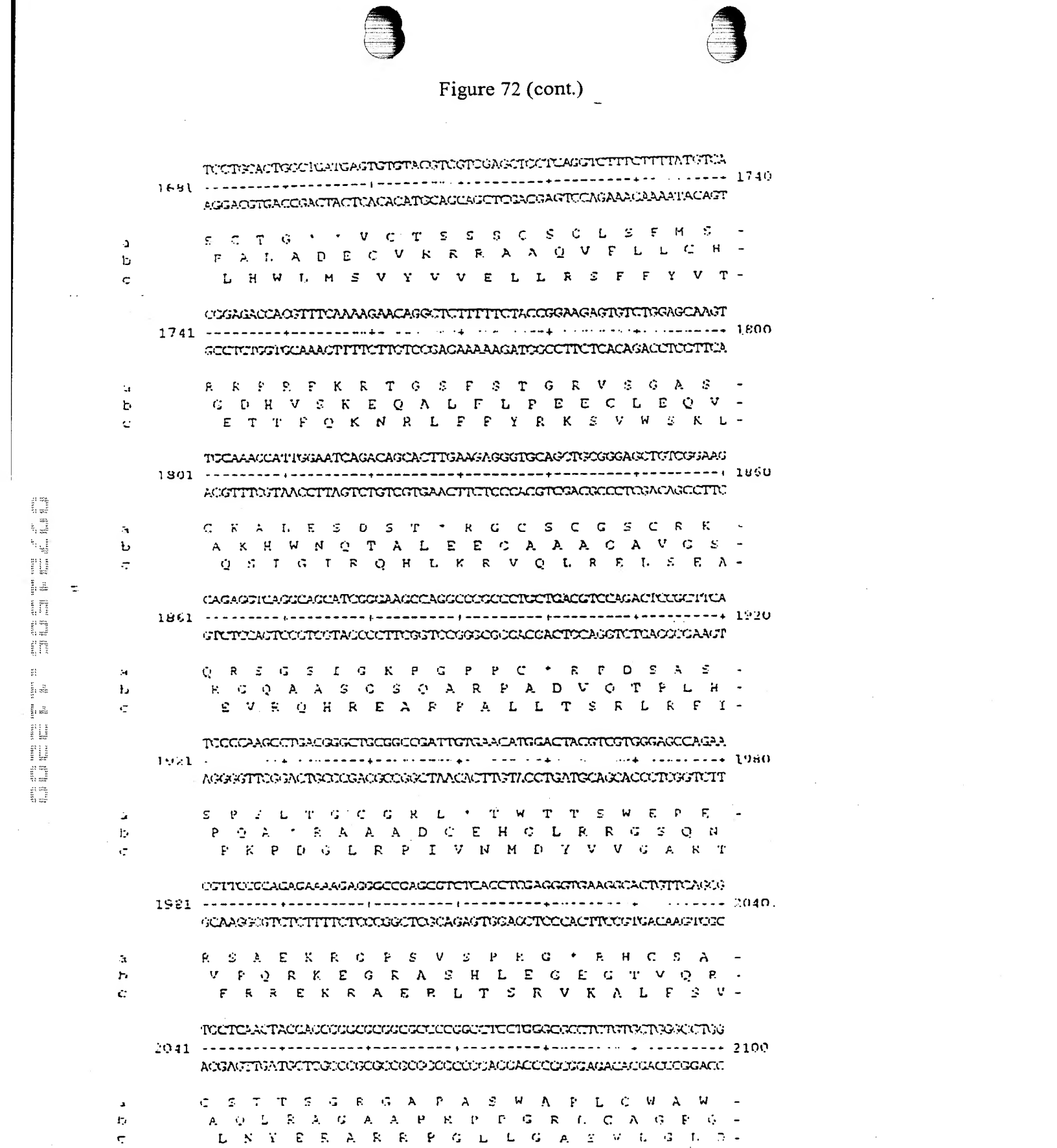


Figure 72 (cont.)

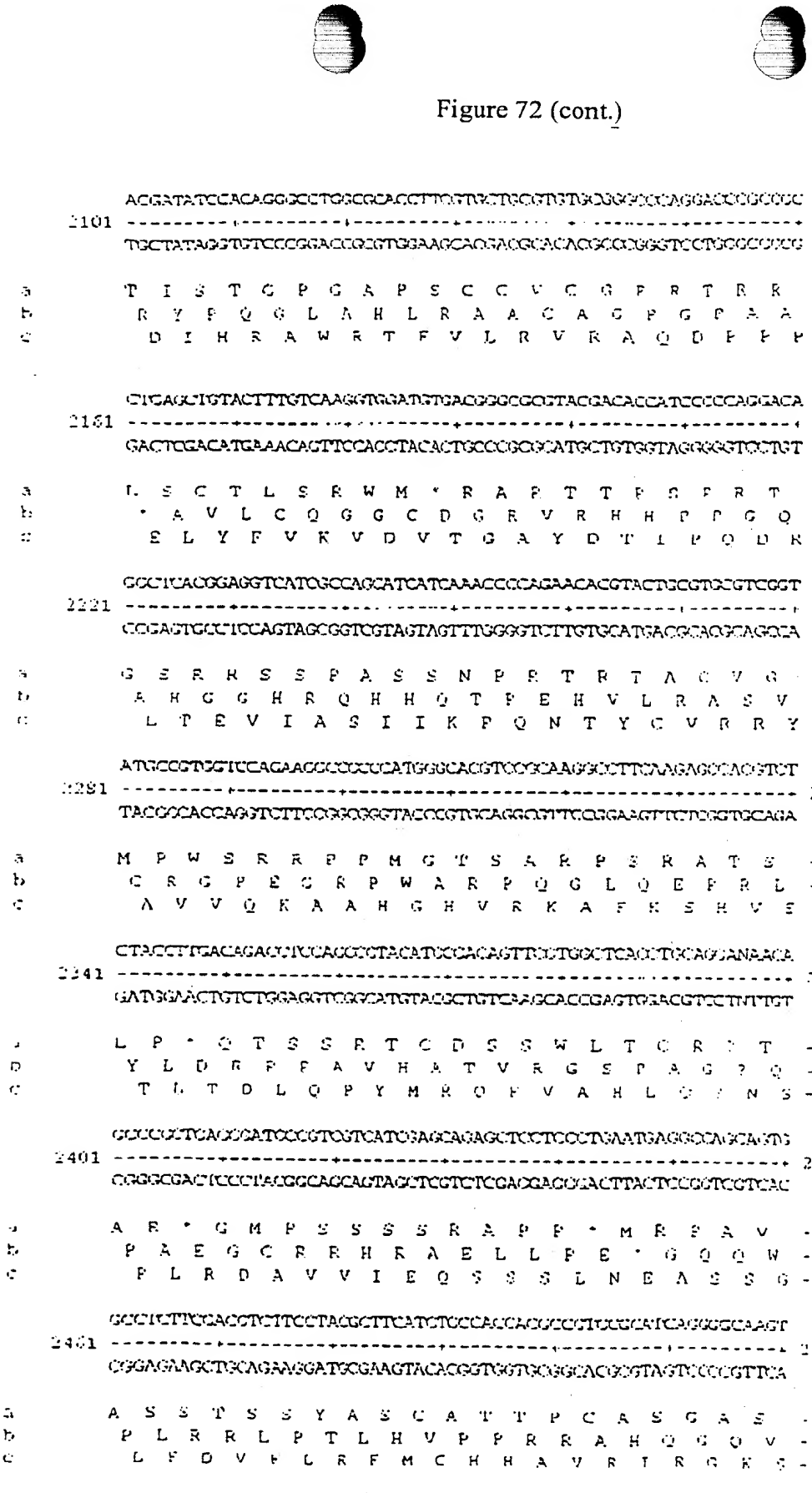


Figure 72 (cont.)

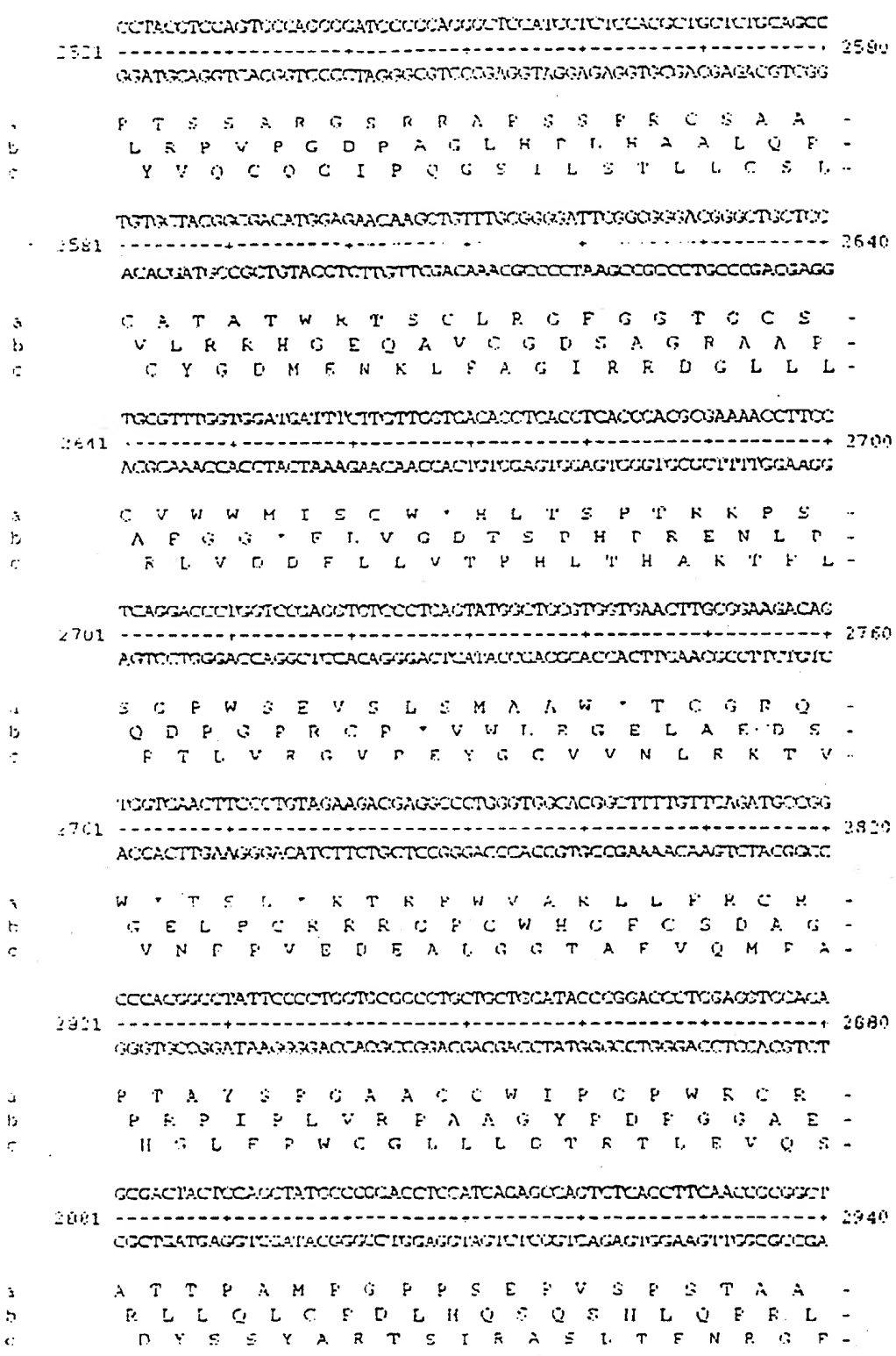


Figure 72 (cont.)

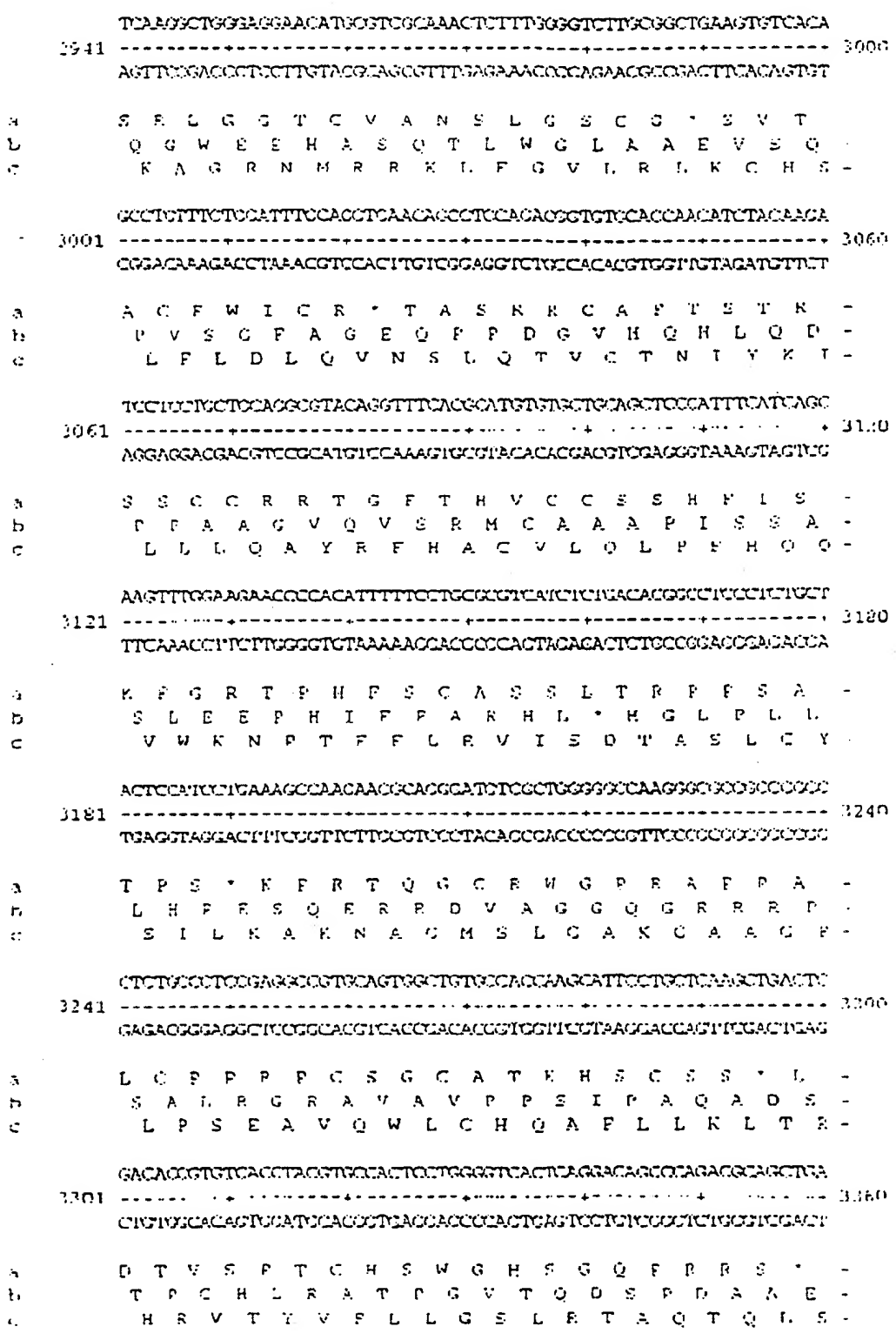


Figure 72 (cont.)

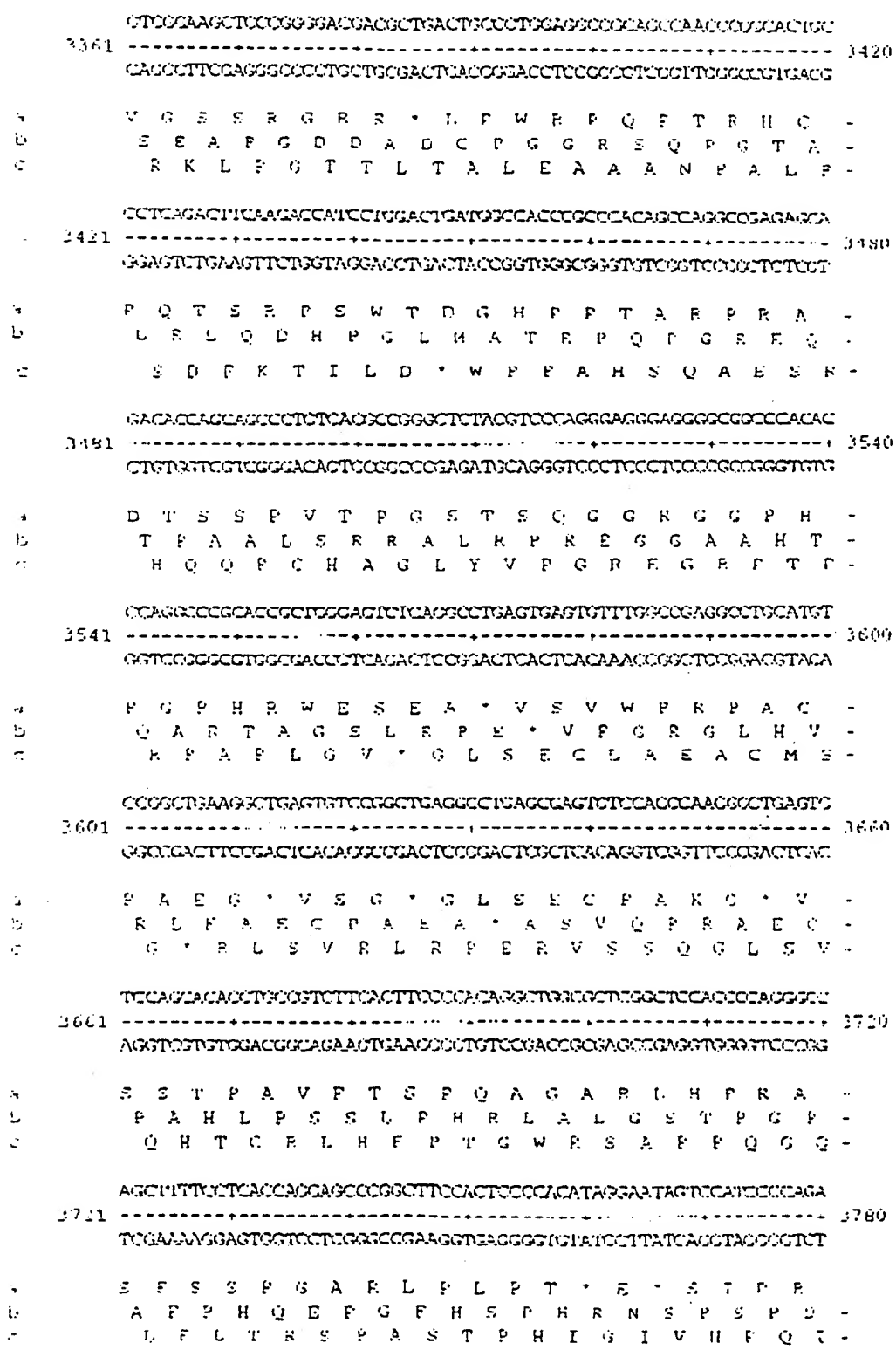


Figure 72 (cont.)

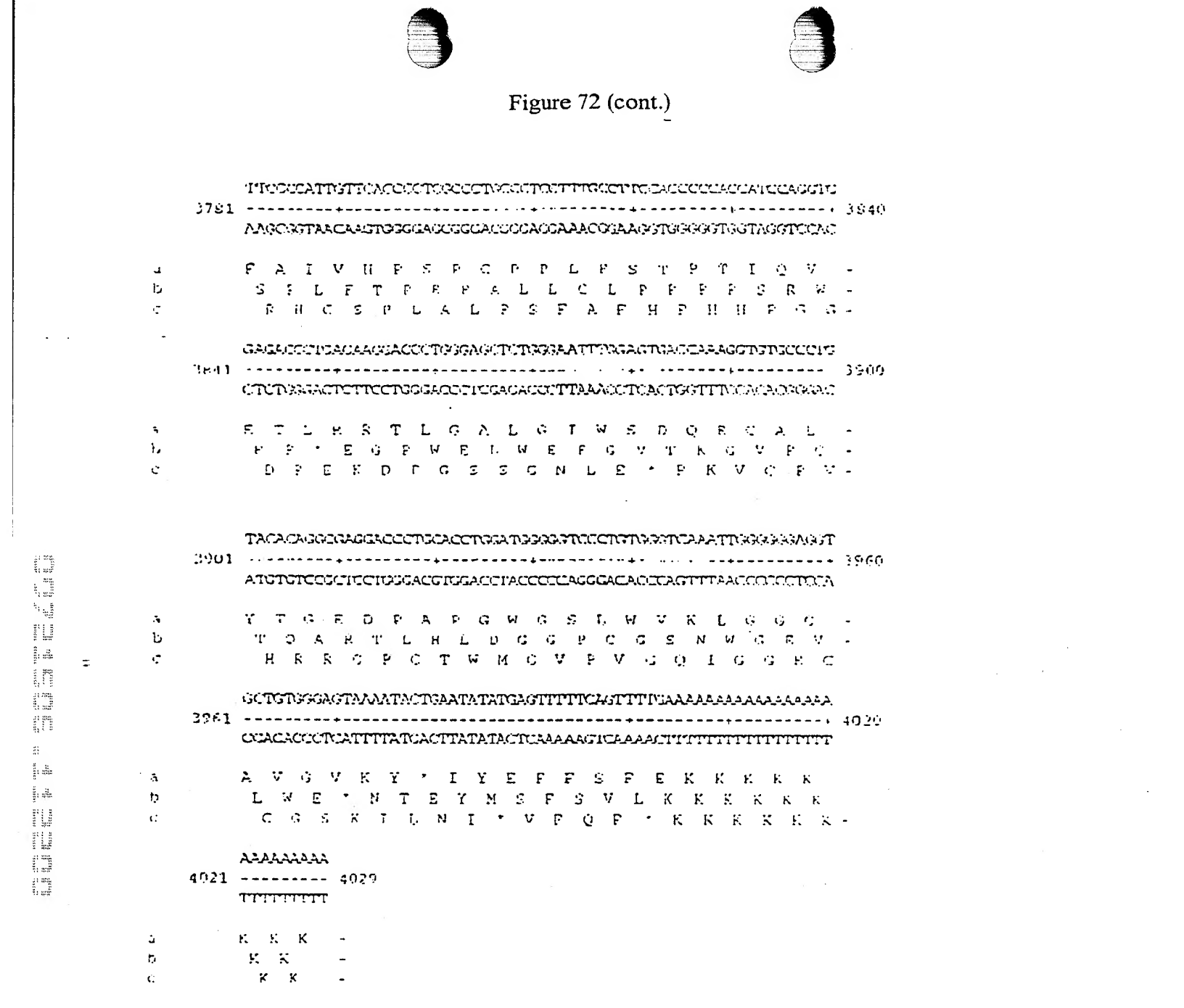


Figure 73

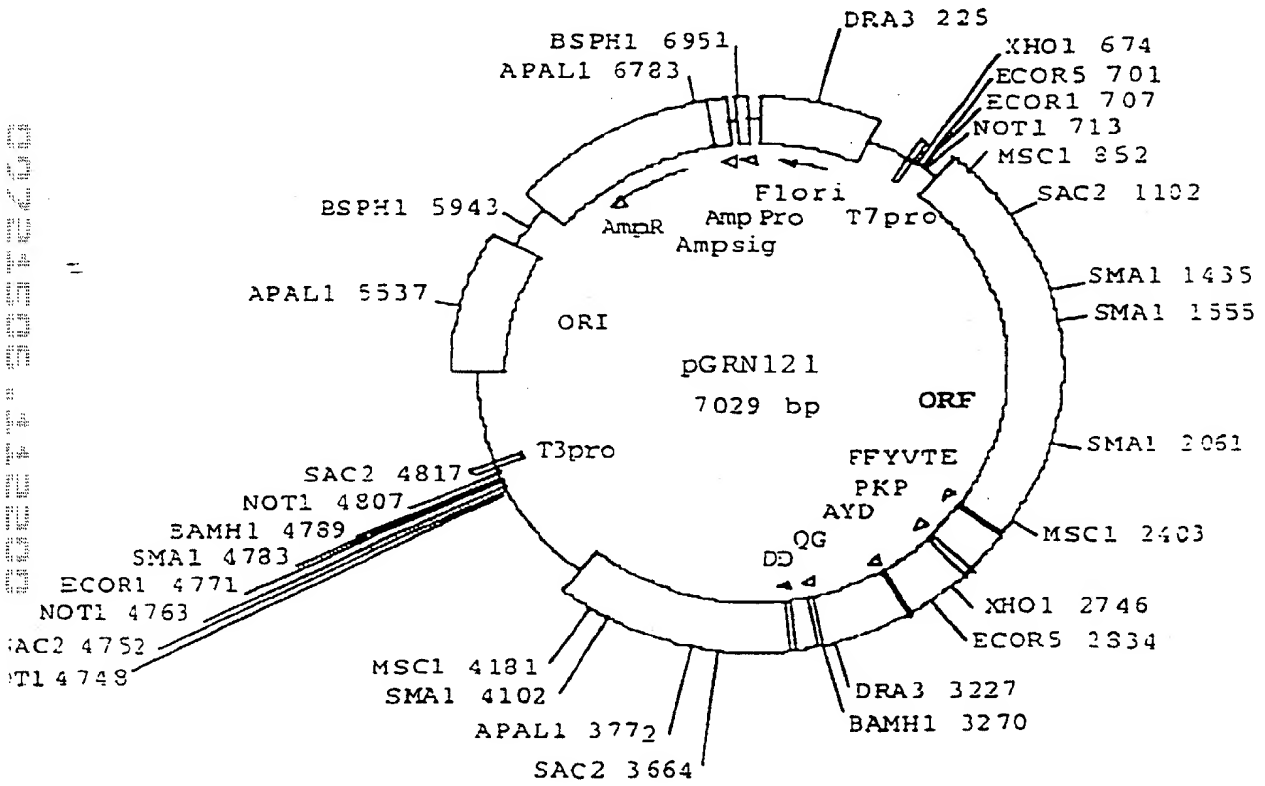

DRA3 225
BSPH1 6951
APAL1 6783
XHO1 674
ECOR5 701
ECOR1 707
NOT1 713
MSC1 852
SAC2 1102
BSPH1 5943
Flori
AmpR
Amp Pro
T7pro
Ampsig
SMA1 1435
SMA1 1555
APAL1 5537
ORI
pGRN121
7029 bp
ORF
SMA1 2061
T3pro
FFYVTE
PKP
AYD
QG
DD
MSC1 2403
SAC2 4817
NOT1 4807
BAMH1 4789
SMA1 4783
ECOR1 4771
NOT1 4763
XHO1 2746
ECOR5 2834
MSC1 4181
SMA1 4102
APAL1 3772
SAC2 3664
DRA3 3227
BAMH1 3270

Figure 74

```
                                                                      1
                                                                     met
GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCCGGCCACCCCCGCG ATG

                                        10
pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

                20                                                  30
his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

                                        40
gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

                50                                                  60
phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

                                        70
ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

                80                                                  90
leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

                                        100
gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG

                110                                                 120
ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC


                                        130
tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG
```

Figure 74 (cont.)

```
                140                                                 150
trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

                                    160
leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

                170                                                 180
ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

                                    190
thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

                200                                                 210
leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

                                    220
val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

                230                                                 240
ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

                                    250
ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

                260                                                 270
ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

                                    280
val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG
```

Figure 74 (cont.)

```
                290                                                 300
gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

                                    310
gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

                320                                                 330
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

                                    340
leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

                350                                                 360
leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

                                    370
glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

                380                                                 390
arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

                                    400
leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

                410                                                 420
val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

                                    430
ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG
```

Figure 74 (cont.)

```
                440                                                 450
ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

                                    460
leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

                470                                                 480
ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

                                    490
his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

                500                                                 510
leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

                                    520
met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

                530                                                 540
gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

                                    550
ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

                560                                                 570
leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

                                    580
arg leu phe phe tyr arg pro ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG CCG AGT GTC TGG AGC AAG TTG CAA AGC

                590                                                 600
ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG
```

Figure 74 (cont.)

```
                                                  610
ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

                620                                               630
leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

                                                  640
pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

                650                                               660
arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

                                                  670
phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG

                680                                               690
gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

                                                  700
thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG

                710                                               720
tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

                                                  730
gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

                740                                               750
asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

                                                  760
his gly his val arg lys ala phe lys ser his val ser thr leu
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC TCT ACC TTG
```

Figure 74 (cont.)

```
                770                                                 780
thr asp leu gln pro tyr met arg gln phe val ala his leu gln
ACA GAC CTC CAG CCG TAC ATG CGA CAG TTC GTG GCT CAC CTG CAG

                                    790
glu thr ser pro leu arg asp ala val val ile glu gln ser ser
GAG ACC AGC CCG CTG AGG GAT GCC GTC GTC ATC GAG CAG AGC TCC

                800                                                 810
ser leu asn glu ala ser ser gly leu phe asp val phe leu arg
TCC CTG AAT GAG GCC AGC AGT GGC CTC TTC GAC GTC TTC CTA CGC

                                    820
phe met cys his his ala val arg ile arg gly lys ser tyr val
TTC ATG TGC CAC CAC GCC GTG CGC ATC AGG GGC AAG TCC TAC GTC

                830                                                 840
gln cys gln gly ile pro gln gly ser ile leu ser thr leu leu
CAG TGC CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC

                                    850
cys ser leu cys tyr gly asp met glu asn lys leu phe ala gly
TGC AGC CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG

                860                                                 870
ile arg arg asp gly leu leu leu arg leu val asp asp phe leu
ATT CGG CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG

                                    880
leu val thr pro his leu thr his ala lys thr phe leu arg thr
TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC

                890                                                 900
leu val arg gly val pro glu tyr gly cys val val asn leu arg
CTG GTC CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG

                                    910
lys thr val val asn phe pro val glu asp glu ala leu gly gly
AAG ACA GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC

                920                                                 930
thr ala phe val gln met pro ala his gly leu phe pro trp cys
ACG GCT TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC
```

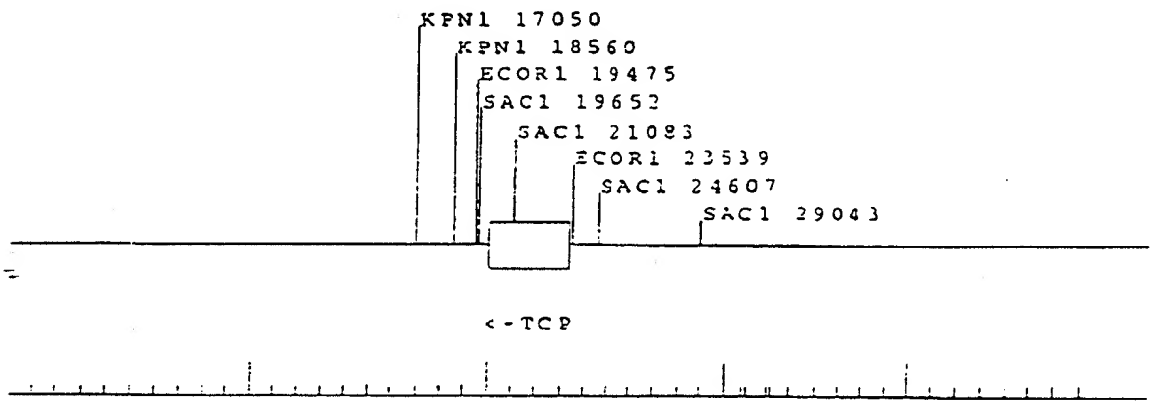
Figure 74 (cont.)

```
                                                940
gly leu leu leu asp thr arg thr leu glu val gln ser asp tyr
GGC CTG CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC

                950                                             960
ser ser tyr ala arg thr ser ile arg ala ser val thr phe asn
TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT GTC ACC TTC AAC

                                                970
arg gly phe lys ala gly arg asn met arg arg lys leu phe gly
CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG

                980                                             990
val leu arg leu lys cys his ser leu phe leu asp leu gln val
GTC TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG

                                                1000
asn ser leu gln thr val cys thr asn ile tyr lys ile leu leu
AAC AGC CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG

                1010                                            1020
leu gln ala tyr arg phe his ala cys val leu gln leu pro phe
CTG CAG GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT

                                                1030
his gln gln val trp lys asn pro thr phe phe leu arg val ile
CAT CAG CAA GTT TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC

                1040                                            1050
ser asp thr ala ser leu cys tyr ser ile leu lys ala lys asn
TCT GAC ACG GCC TCC CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC

                                                1060
ala gly met ser leu gly ala lys gly ala ala gly pro leu pro
GCA GGG ATG TCG CTG GGG GCC AAG GGC GCC GCC GGC CCT CTG CCC

                1070                                            1080
ser glu ala val gln trp leu cys his gln ala phe leu leu lys
TCC GAG GCC GTG CAG TGG CTG TGC CAC CAA GCA TTC CTG CTC AAG

                                                1090
leu thr arg his arg val thr tyr val pro leu leu gly ser leu
CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC
```

Figure 74 (cont.)

```
                1100                                                1110
arg thr ala gln thr gln leu ser arg lys leu pro gly thr thr
AGG ACA GCC CAG ACG CAG CTG AGT CGG AAG CTC CCG GGG ACG ACG

                                        1120
leu thr ala leu glu ala ala ala asn pro ala leu pro ser asp
CTG ACT GCC CTG GAG GCC GCA GCC AAC CCG GCA CTG CCC TCA GAC

                1130    1132
phe lys thr ile leu asp OP
TTC AAG ACC ATC CTG GAC TGA TGGCCACCCGCCCACAGCCAGGCCGAGAGCAGA

CACCAGCAGCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGCCCACACCC

AGGCCCGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCC

GGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTC

CAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCCAGGGCCAG

CTTTTCYTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATCCCCAGATT

CGCCATTGTTCACCCYTCGCCCTGCCYTCCTTTGCCTTCCACCCCCACCATCCAGGTGGA

GACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTA

CACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGC

TGTGGGAGTAAAATACTGAATATATGAGTTTTTCAGTTTTGRAAAAAAAAAAAAAAAAAA

AAAAAAAAAA
```

Figure 75


KPN1 17050
KPN1 18560
ECOR1 19475
SAC1 19652
SAC1 21083
ECOR1 23539
SAC1 24607
SAC1 29043
<-TCP